IKED

INTERNATIONAL ORGANISATION FOR
KNOWLEDGE ECONOMY AND ENTERPRISE DEVELOPMENT

GLOBAL IDENTITY NETWORKING OF INDIVIDUALS

The Individualised Digital Identity Model

A User-centric Framework of Identity Management Services

Report from the GINI Consortium

October 12, 2011

Project Name	GLOBAL IDENTITY NETWORKING OF INDIVIDUALS
Project Number	FP-258630
Work Package	WP1: The GINI Conceptual Framework
Document title	A User-centric Framework of Identity Management Services
Document type	Report
Deliverable number	D1.1
Editors	Thomas Andersson, Lefteris Leontaridis (IKED)
Date	2011.10.12
Version	1.1
Status	Final
Total number of pages	90

Table of Contents

1 Introduction

GLOBAL IDENTITY NETWORKING OF INDIVIDUALS (GINI-SA) works towards the vision of a Personal Identity Management environment where individuals will be able to manage their own identity space (INdividual Digital Identity – INDI), where User-specific identity attributes and related services will be available in the INDI environment which is also marked by the development of demand-driven protection and privacy enhancement provisions.

As of today, we are far from a situation where individuals have sufficient control of their identity attributes. Some twenty years ago, when the web had hardly come off the ground, digital communications were not associated with any issues in regard to trust, basically as they then played merely a complementary role back-to-back with tangible real-world interfaces between people who already knew each other. As the Internet has evolved, and more recently also other forms of social networks and diverse mediums for digital communication and exchanges, into a vast but unwieldy universe of intensive free-standing interactions, users are mostly lacking the information as well as the tools to take action in support of authentication and orderly management of their identities. There is a lack of mechanisms to allow for the emergence of agents, or operators, capable of responding to the needs of identity management, by developing an array of services that would be relevant for responding to such demands.

Unless these issues are addressed effectively, we are heading for a much more chaotic situation than what we have today. That would be a future marked by high uncertainty whom (or what) is "at the other end of the line", with what rights and obligations, what data there is about us, by whom and how it is used, and so on. It would be a future marked by lack of choice in managing an identity space, in which users and individuals would be the easy prey of identity theft and other forms of data misuse.

Another future ought to be possible. This would be one in which there would be an orderly response. Operators would be established to develop differentiated user-driven services in support of INDI, which would also be addressing associated challenges with regard to, e.g., privacy, security, accountability and trust. An important question is what with what it takes to get there.

Many countries are faced with similar challenges in this respect. Whereas the search for solutions is underway in various parts of the world, however, there are marked differences in views and favoured approaches. It is still not clear what precise role governments should play, or how other stakeholders should be engaged. The National Strategy for Trusted Identities in Cyberspace, launched by the United States, carries strong elements of central coordination. The Japanese Government, on the other hand, has taken the lead in engaging private industry on these issues to work out a "new Internet". In Europe, there is a strong belief in the need of effective collaboration in establishing a comprehensive framework for identity management which is capable of responding to the various needs at the local, national, and international/global levels.

The GINI project, which is undertaken as a horizontal action and run in close collaboration with the European Commission, reflects that conviction. This document provides an assessment of existing and emerging technical solutions and standards that relate to the management and consumption of individual digital identities. It initiates a systematic architectural approach to the INDI characterisation, focusing on the high-level conceptual architecture motivations, requirements and constraints.

1.1 Conventions and Definitions

This document introduces the terms related to the management and consumption of digital identities that are of relevance for the conceptualization of the GINI vision. While these terms are

introduced within the context of broader conceptual and practical aspects of digital identities throughout the text, an additional glossary is provided for giving precise definitions of these terms.

Throughout the text, existing standards, products and research projects related to digital identities are taken as a baseline for the conceptualization of the GINI vision. In order for the document to stay focused, descriptions of standards and projects are kept to a minimum within the text. For readers who want to learn more about a certain standard or product, a full list of all referred to standards and projects is given in annex A to this document.

This document presents the state of the art in managing and processing digital identities in order to point out how GINI will go beyond this in order to empower citizens to gain control over their digital identity. Within the text, the respective GINI objectives and requirements are highlighted by a shaded frame.

1.2 Organisation of this Document

This document introduces a user-centric framework for identity management and consumption. In its first chapters, the core concepts, patterns, and technologies are briefly presented and revised to prepare for the subsequent documents. Chapters 2, 3, and 4 provide information on the potential interplay of the different GINI actors and entities, in addition to being anchors for the more detailed and comprehensive work ensuing in WP2, WP3, and WP4. The heart of this document is **Chapter 5,** in which the INDI ecosystem and vision is outlined and filled with real life use-cases and application scenarios. Furthermore, Chapter 5 serves to unite the provisions brought forward in the previous chapters, and puts the issues associated with technology into a broader perspective. A Glossary is annexed which reflects terminology both already current as well as specific to GINI.

Potential shortcomings of current technology, and other gaps which require extensive additional research, are only briefly mentioned in this document but discussed in detail in subsequent work (notably D2.2).

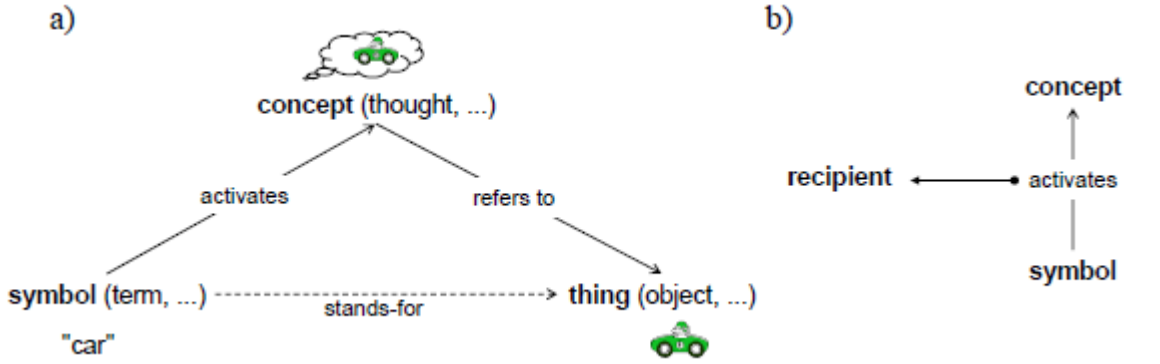
2 Digital Identities

The objective of this section is to set the scene for the various flavours digital identities have: from the perspective of underlying concepts as well as their application areas. As a stock taking exercise we give a crisp glance on the various facets, referring to relevant literature for further details. The main purpose here is to provide readers either not being experts in field, or looking at digital identity from a particular angle, with a succinct overview of the various different, perhaps even contradictory, existing paradigms. This shall prepare for the GINI vision considered at the end of this document and prepare the reader for the more in-depth deliverables on the architecture, technology and other aspects addressed in the subsequent documents.

2.1 Principals, Subjects and Identities

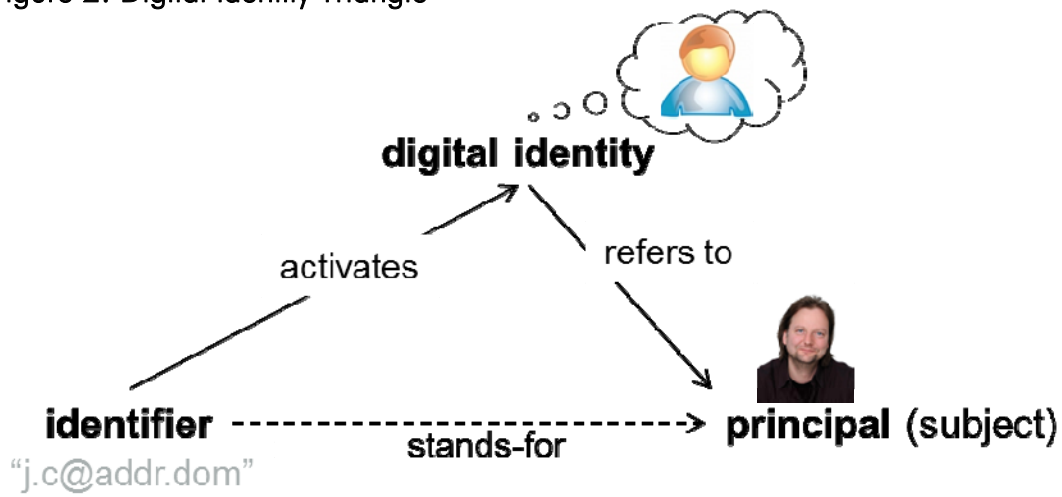
The relevance of digital identities for any digital communication can best be expressed in terms of the well known semiotic triangle. In this model, a symbol is a syntactical entity that activates a concept or thought at the recipient of this symbol - the meaning of the symbol. The concept refers to the real-world object.

Figure 1: Semiotic Triangle



By exchanging "object" for "principal", "symbol" for "digital identifier" and "concept" for "digital identity" we get: A digital identifier is a syntactical entity that instantiates a digital identity for a subject. Therefore, whenever two parties have to share a perception of the same person or system object they make use of a digital identity that represents this real-world entity. Given this model, a digital identity can be considered a projection of a real-world entity onto a set of identity claims within a virtual space.

Figure 2: Digital Identity Triangle



A principal is a real-world entity; its virtual counterpart is often called a "subject" – e.g. a computer process running on behalf of the principal. A real-world entity sitting in front of his computer and at the same time controlling processes in the virtual world of that computer, is therefore at this same time a principal (for someone who observes him from within the real world) and a subject (for his computer who just "sees" him through some signals sent from an interface device).

To avoid too much confusion on terms throughout this deliverable, the term "principal" will only be used if we explicitly want to refer to a real-world entity, in all other cases the more common term of a "subject" will be preferred.

2.2 The Need for Digital Identities

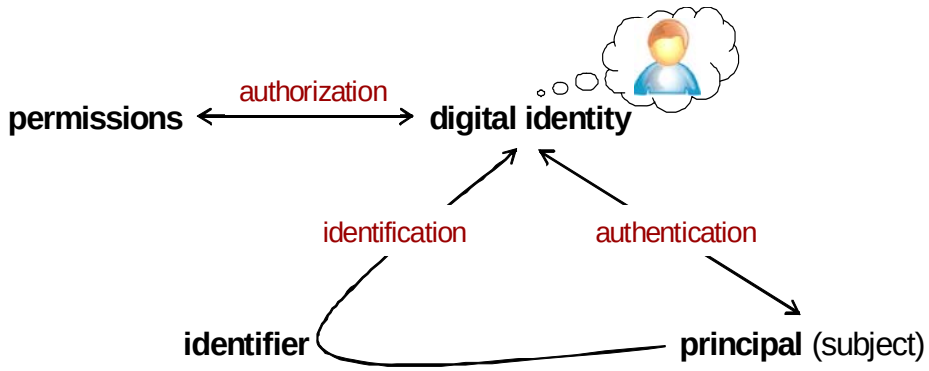
The term "identity" plays an important role in real life as well as in the digital world. As elaborated in the previous section, the main purpose of an identity is to act as a concept or thought that refers to a real-world entity and its specific characteristics. Whenever someone enters the US, an identifier of this person (e.g. a passport number) is used to activate the digital identity of the citizen that exists within the databases of the US Department of Homeland Security. Attributes of this digital identity within the database are projected on the immigration officer who then either accepts or denies immigration.

This example shows a common characteristics of digital identity use cases: claims and properties of a digital identity control the privileges of a real-world entity. These privileges are controlled by associating permissions with either the digital identity as a whole, or with certain properties of the digital identity. In the domain of identity management, binding (assigning) permissions to an identity or an identity attribute is usually referred to as "authorization[1]". In order to enforce the authorization assigned to a digital identity, this digital identity must first be activated through "identification" of the acting principal. An additional verification that the activated identity really refers to the identified principal can be obtained through "authentication" of the principal.

[1] Not to be mistaken for the authorisation in reference to access control (enforcement).

In the following sections, the concepts of identification, authentication and authorization are covered in more detail.

Figure 3: Identification, Authentication and Authorization



2.2.1 Identification

Identification defines the association between a personal characteristic and a subject representing various attributes. A characteristic that allows for an unequivocal identification within a closed context is called an identifier. For example, identification can be described as the association between a person (subject) and the full name. In this case the name e.g. "John Doe" identifies the person "John Doe".

An identification process defines the presentation of an attribute a person or object can be uniquely identified with in a given context. According to the example above, the person "John Doe" can be identified by presenting his name. Unique identification is only possible as long as no other person with the name "John Doe" exists within the given (closed) context. Otherwise additional attributes for unique identification would be required.

Unique identifiers should only be assigned to a principal by an entity that is authoritative for the particular domain. Each subject identifier originates from an assigning authority. The value of this subject identifier must be unique within the concrete application domain. During the process of identification, the assigning authority can either be implied or must be explicitly stated (e.g. by using an identifier that univocally identifies an assigning authority. This identifier has an assigning authority, too, which is either implied or defined within the context of the identification process.)

GINI Objective: Digital identities are often closely linked to identifiers: depending on the identifier used a certain digital identity is activated. It is a goal of GINI to decouple the activation of digital identities from the use of any particular identifier, and to support the use of multiple identities and/or identifiers.

Table 1 lists some commonly used identifiers and their respective assigning authorities.

Table 1: Commonly used Identifiers

Identifier	Assigning Authority	Remarks
User ID	Domain Administrator	Examples of domain administrators are a company's system administrator, an application's administrator, etc.

Passport Number	Ministry of the Interior	
Social Security Number		
e-Mail Address		
X.500 Subject Name		
X.509 Certificate Number	Accredited certificate issuer (certificate authority)	
URL	Domain Administrator; usually confirmed by an Identity Provider	*Some words on URL-based ID (openID etc.)*

2.2.2 Authentication

The term "authentication" defines the process of verifying a subject's identity or other claim, e. g. one or more attributes. The mere assertion of (an identifier associated with) a digital identity depicts only a claim. In many situations additional proof will be required in order to corroborate that the subject is in fact who it claims to be. Such proof is provided during the authentication process. By checking this proof, a computer application can make assumptions about the accuracy and trustworthiness of the presented claim. With respect to authentication two important aspects must be considered:

- a computer application only "knows" about digital identities. It has no way to process real world objects beyond their respective digital representations. Authentication only proves that there is some real world entity that corresponds to the digital identity.
- Authentication significantly increases the likelihood that an active digital identity corresponds to an identified subject. However, no authentication protocol can provide 100% assurance.

Authentication of humans is based on one or more of the following factors:

- something you know; e.g. a password or a personal identification number (PIN);
- something you have; e.g. a smart card with an embedded secret, or a (physical or digital) key;
- something you are; a physical characteristic such as a fingerprint, iris scan or voice pattern.

Authentication is usually stronger with a higher number of factors and indicators to be applied. For accessing resources with high or even very high requirements on security and privacy, two-factor authentication is sometimes used (e.g. a key you have is released with a PIN you know). Additionally, the authentication means, processes, and potential supportive indicators are qualitatively assessed and grouped based on their individual stability and confidence. The result is the assignment of a certain category that is clearly stating the authentication assurance of the current authentication scheme. A typical state-of-the-art example is the four levels of assurance as de-

scribed in ISO/IEC 29115 / ITU-T X.eaa Entity Authentication Assurance Framework and its concrete application in STORK[2].

2.2.3 Authorization

Authorization is essential to control access to protected resources. Through authorization, (access) rights are assigned to a digital identity. This assignment of rights can either target a single identity directly or make use of higher order identities, such as groups or roles. Examples of such indirections are:

- in a hospital access to a radiologic system is only granted for employees that belong to a group of people that is approved by the head of the radiologic department (rights are bound to groups of individuals);
- an online book store only allows people to place orders who are at least of age 18 (rights are bound to attributes, e. g. the age of an individual).

Authorization is typically preceded by authentication (see next section for exceptional cases).

Authorization may follow different purposes. The most prominent are:

- Put the owner of a resource in full control of who is allowed to access the resource.
- Protect the confidentiality or integrity of resources by matching protection demands with pre-defined levels that are assigned to certain groups or roles within an enterprise.
- Align permissions with an enterprise's organisation ("need-to-know principle").

2.2.4 Identification without Authentication

In many access control scenarios a subject is processing data that is linked to a different individual than the subject (e. g. a physician processing a patient's identity data). In these scenarios usually only one of the individuals is authenticated while the other individual is only identified. A typical example for this is the sharing of health data:

- A physician needs to access health data about a patient. She identifies and authenticates herself towards her local system. She then identifies the patient and sends a data access request to a health record system. The health record system verifies whether the physician is authorized to access the identified patients' health data. If this check succeeds, the requested data is released to the physician.
- A patient wants to release a lab report in his personal health record to a physician. He identifies and authenticates himself towards his personal health record application. He identifies the physician and obtains the physician's contact data from the physician's digi-

[2] STORK, D2.1 – Framework Mapping of Technical/Organizational Issues to a Quality Scheme, http://www.epractice.eu/en/library/292295

tal identity data. Using this contact data the patient transmits the lab report to the physician.

These examples even show the strong dependency between authentication and authorization: Authentication is only applied to the actor whose privileges are to be verified. For the other actor no authentication is required as he is the passive part (target) of the access control policy that is to be enforced.

A specific challenge for such scenarios is that the subject (active part of the access control scenario) may not be aware of a unique identifier of the individual he needs to identify. In these cases means must be provided to allow for identifying individuals either in an interactive manner and/or by querying against a set of identity attributes that are each not unique when taken in isolation (e.g. demographic data).

GINI Objective: GINI must allow for identification even in scenarios where no unique identifier is available, subject to applicable policy. As this process may disclose protected identity data the citizen must be in full control about who is allowed to investigate on his identity and what identity data and identification processes may be used for this kind of identification.

2.2.5 Authentication without Identification

Identity related user information like e.g., date of birth, name, and so forth becomes increasingly accessible in the web-based ecosystem, which puts a growing demand on the challenging task of protecting the user's privacy. To solve this problem, the user should be allowed to control the dissemination of identity attributes and to access services based only on attributes necessary for service usage, instead of a larger set of identity related information needed for authentication. The best-known ideas for such control are group signatures and anonymous credentials.

A group signature (or anonymous credential (AC)) is a set of identity attributes certified by e.g., Identity Service Provider, Government or any other trusted certification authority. With the help of ACs the disclosure of user information not necessarily needed for the consumption of an individual service but usually collected during an authentication process could be prevented. Access management systems usually collect a larger set of data that they use for authentication purposes and some of them even use digital certificates to further verify the identity of a user. Based on the kind of certificate or the set of data to be used much more data is collected the actually needed for a specific service, which makes such systems an interesting source of user related information. In the case of fraudulent access to the system it could lead to a breach of sensitive information, some of which could be used later by an attacker.

The following scenario should help to illustrate the above. One evening Mike decides to go to a casino for a game of poker and some drinks. Reaching the Casino, he is asked by the Croupier for his driver's license to prove that he is of legal age to play poker and drink. By handing over the driver's license to the croupier, Mike is transferring more attributes then actually needed to prove his age. In a face-to-face situation this may not be a big issue. But in environments such as the Web, an entire digital certificate, which would be the digital representation of the driver's licence, may be exposed to the whole world over the Net, where its contents can be sniffed and stolen by attackers interested in stealing authentication credentials. It is also very common in today's times of cheap storage space that information may be stored indefinitely. Minimal disclosure reduces that problem by only providing the necessary information needed to grant access to a service. The entire set of the user's attributes or credentials does not need to be disclosed to the service-requesting authentication.

An anonymous credential-aware ecosystem consists of users and organisations, which know the user only by pseudonyms. A user can create pseudonyms to be used in certain contexts or scenar-

ios e.g., Mike: for Casino and Beer, Paul for opera and theatre, these pseudonyms of the same user cannot be linked by the organisation or by other users in that ecosystem. This means an organisation can issue a credential to a pseudonym (Mike), and the corresponding user can prove possession of this credential to another organisation, which knows him by a different pseudonym (Paul), without revealing anything more than the fact that he owns such a credential. Credentials can be for unlimited use (multiple-show credentials (Idemix)) or for one-time use (one-show credentials (U-Prove)).

2.3 Identity Lifecycle

Understanding the lifecycle of digital identities facilitates Identity Management (IDM) in many respects. It allows for mapping identity-related operations to their corresponding lifecycle phases.

Figure 4: Identity Management Lifecycle

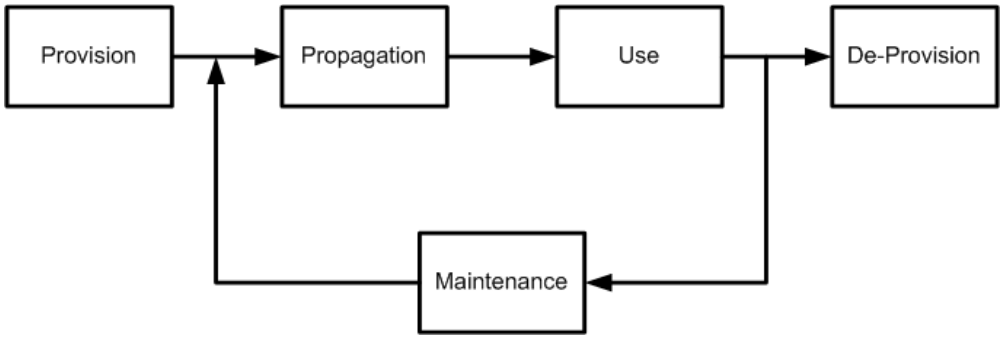


Figure 4 shows the IdM lifecycle phases as described by Windley in "Digital Identity"[3]. The identity lifecycle starts with provisioning, i.e. the creation of a digital identity. The identity is then propagated to the places where it is used. Identity data may change over time or additional attributes may be added to a digital identity. During its lifetime, a digital identity is usually subject to a continuous maintenance process. Finally, if a digital identity is not needed any longer, it is deprovisioned or revoked. We give a brief overview of each of these phases in the next subsections.

GINI Objective: Individuals shall have as much control as possible over every step of his digital identities' life cycle.

2.3.1 Provisioning

Provisioning of a digital identity is the process whereby the digital identity record is created. This digital identity record contains the attributes that are initially associated with the identity. This process can take place in an entirely automated fashion or may be carried out manually. In case of manual instantiation, identity data can either be provided by case officers, agents, administrators or by self-registration. The latter case is typical in today's internet applications like web mail providers, online shops, discussion forums, chat portals, etc. Many Web applications allow for self-

[3] Phillip J. Windley. Digital Identity. O'Reilly.

assertion and do not rely on verified identity data. Users can enter their name, age, date of birth, e-mail addresses or postal addresses into web-forms. This data is then taken to create the identity record and to provide a digital identity for the system.

2.3.2 Propagation

After creation, a digital identity is usually propagated so that other system components and people can access and use it. This is not the case when the identity is just used locally and stored as file or database record. However, in more complex identity management systems, identity data may be communicated to (multiple) central directory services, meta-directories or other places. If this propagation also concerns external entities across organisational boundaries we refer to a federated propagation.

Errors during the propagation process may lead to inconsistent states across the system. Since propagation involves multiple operations, appropriate safeguards must be foreseen. IDM systems therefore often implement a transaction workflow process managing for propagating a digital identity. Propagation of identity data is a crucial process and if not correctly handled may threaten the subject's privacy.

2.3.3 Use

In this phase certain identity attributes are used to carry out particular operations related to the digital identity. The "use phase" is the most visible phase of a digital identity. System components and people use it to authenticate and authorize resources, e.g. for communication or transactional operations. Systems that make use of digital identities are commonly called "identity consumers" whereas the system that prepares identity data for use is called the "identity provider".

2.3.4 Maintenance

Identity data is usually not static and can/must be changed over time, e.g. if people change department or roles or if users change their home address, etc. In this case maintenance operations force the identity management lifecycle to continue the workflow with the propagation phase. Changed identity attributes must be propagated to the system components and people that use it. An efficient identity maintenance process can significantly reduce costs, e.g. if users are encouraged to change their data by themselves or by the provision of auto-recovery mechanisms (in case of lost passwords, etc.).

2.3.5 De-Provisioning

The de-provisioning of digital identity data is as important as the provisioning. It can be seen as the reverse process of identity provision and propagation. In particular, in a federated scenario or a direct trust relationship, the deleted data must also be propagated to the affected entities (directories, etc.). If de-provision is not correctly handled, this may affect the confidence in the security of a system. Consider the typical set-up of an enterprise comprised of many separate systems and applications. Upon the registration, provisioning, and propagation of a new digital identity, all affected systems and applications are informed about the new identity. This may well include third parties, such as a payroll or VPN-access service. An employee leaving a company, whose digital identity (user account, privileges, payroll messages, VPN credentials, etc.) is not fully deleted may cause an unauthorized access or provide an additional attack vector for hackers, since unused accounts are usually not that closely monitored as used ones. Incorrectly de-provisioned

data may also lead to confusion (wrong user statistics), compliance issues in sensitive domains or to additional licensing costs in case of a per-person licensing model.

2.4 Identity Types

Digital identities have many flavours. Identities can be distinguished e.g. between the location where they are created and used, where they originate and where they are transferred, etc. We give a short overview on identity definitions found in today's IdM systems.

2.4.1 Partial Identities

As digital identities often map onto a real-life entity, they might disclose information about this entity to all parties that participate in the digital identity lifecycle. For individuals as such entities, there is a strong demand to control the identity information that is stored and released for use in order to protect the individual's privacy. Controlling the disclosure of identity information imposes restrictions on what identity data is disclosed to whom for which purpose.

Subsets of an individual's full identity record are called "partial identities". Based on the model introduced in section 2.2, a partial identity can be considered a partial projection of a real-world entity onto a set of identity claims. A partial identity usually is limited to the identity claims an identity consumer needs to perform its intended transaction. The derivation of this proper subset of the full digital identity can be done at different steps during the identity lifecycle:

- Provisioning: This kind of a partial identity is usually called an account – a self-contained set of identity information that is used for a specific purpose only and that is maintained for serving this purpose only. E.g., when setting up an account with an Internet access provider, an individual will only provide information that is required for authentication, non-repudiation, and billing. Other subsets of the full set of identity attributes will be disclosed for participation in social networks.

- Propagation: In instances in which a large set of identity attributes is maintained within an Identity Management system, only subsets of these attributes are propagated to systems and applications that perform access control decisions based on these attributes.

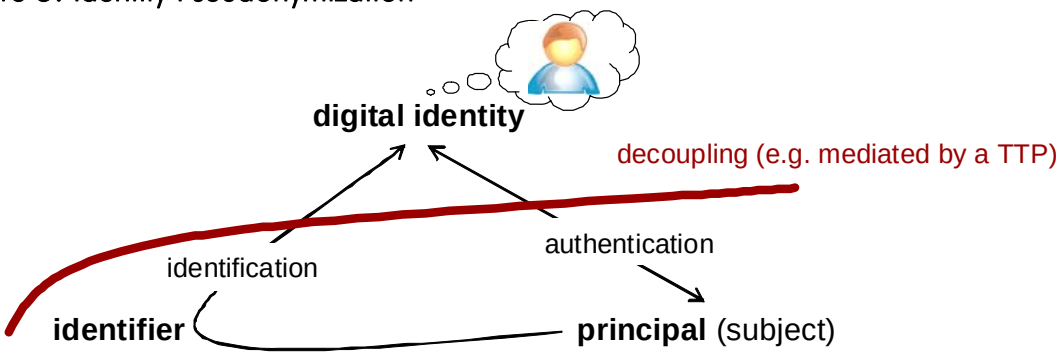
- Use: During propagation only a minimum set of attributes is provided to identity consumers. Upon use of the identity the consumer requests required identity information from a dedicated identity source (e.g. a system which stores the complete set of identity data available within that system). The identity source only provides the information that is needed for the identity consumer to the purpose of the transaction.

GINI Objective: The separation of identifiers and other identity attributes enables systems to minimize the disclosure of private identity data. Due to the on-demand provisioning of identity data the provided set of identity data can be determined at usage time and therefore can be adapted to the minimum requirements of the identity consumer. GINI will provide means on how this partitioning of identity data during the use-step of the identity life cycle can be implemented in a user-friendly manner that allows the user to maintain control with minimum administrative effort.

2.4.2 Pseudonymous Identity

Pseudonymization is a powerful means for protecting the privacy of individuals. Pseudonymization decouples a digital identity from the real-world entity but preserves the univocal linkage to a unique entity. The only difference is that an identity consumer does not directly interact with the individual and is not provided with information that enables her to ascertain the real-world identity of the subject.

Figure 5: Identity Pseudonymization



From an identity consumer's perspective, a pseudonymous digital identity can be treated as a "real" digital identity. Due to their univocal – yet secret – linkage to a real-world entity, pseudonymous identities can be used to set up accounts and profiles (e.g. for managing preferences).

Reversible pseudonymization requires the use of a trusted third party as a mediator that decouples a digital identity from the underlying real-world entity. The mediator is responsible for mapping identifiers and for maintaining the correct linkage between the respective digital identities. The trusted third party is able to disclose the relationship between a pseudonymous identity and a real-world entity upon request, subject to applicable policy. This property is e.g. used in conjunction with medical studies: A patient can provide medical data for a study and remain pseudonymous; that patient's entire data is linked to the same pseudonym but the study manager does not know the real name of the patient. In case that the processing of the patient's medical data yields to information that requests immediate medical action on the patient, the trusted third party is able to get in contact with the patient and to inform him about the study findings.

2.4.3 Anonymous Identity

Identities may be completely anonymous in a way that it is impossible to link a digital identity with an identifiable subject. An anonymous identity can be considered as a partial identity whereby the available identity claims are not sufficient to derive an identifier or otherwise link to any real-word entity. Usually anonymous identities are temporary and only used for a single transaction.

Within this single transaction an anonymous identity rushes through the whole identity lifecycle: it is propagated to its consumer, processed and then forgotten. A typical example for this is the collection of statistical data in an online questionnaire: A user fills in some web forms; the data is submitted to a web portal, processed and stored in a database. After this the session is closed and all that remain is a dataset in a database that is not linked with any digital identity.

Due to the data tracks that are created even through the short life cycle of an otherwise anonymous identity, anonymity is hard to preserve. Other threats against anonymity include the linkage of identity claims with outer identity data sources (e.g. social networks) in order to derive a set of

candidate persons that can be further restricted by analysing even more identity data sources. Therefore, most anonymous identities are rather pseudo-anonymous: viewed in isolation they do not disclose a real-world entity, but if life cycle data tracks and public identity data sources are analysed one may be able to unveil the principal they referred to.

2.5 Identity Domains

The validity (and context) of a digital identity is usually bound to a certain domain of applicability. The linkage between the identity and the underlying real-world entity may only be fully resolved within this particular application domain. Outside the validity domain a digital identity may only represent a set of claims that cannot be adequately associated with any semantics (referring to section 2.2: identity consumers have no "thought" about this identity).

2.5.1 Local Identity

A local identity can be seen as a digital identity that is created and used only in a closed environment or domain. This kind of identity is usually created and maintained by the user or the system. A typical example is a local password-based access, where user accounts, associated groups and passwords are stored within a file in the host environment.

2.5.2 Global Identity

Like a passport in the real world, a global identity (GI) serves to identify entities in a broader context, i.e. across local domains or within one global computing ICT infrastructure, e.g. the Internet, the web or a grid structure. GI is thus an interesting application area for e-Government, but as well for private sector applications, e.g. in the context of a virtual marketplace for e-Commerce. Besides obvious benefits, such a model raises privacy and security concerns as well as the threat of identity theft with broader impact. Other challenges concern responsibilities for provisioning, maintenance and de-provisioning of identities when applied globally across different autonomous sectors.

Global identity Management (GIM) must not necessarily be realized by introducing a single global IMS. GIM can also be realized by interconnecting different global IMS and ICT infrastructures. A framework for a federated GIM is introduced and discussed by Mehrdad et al[4]. Technologies for associating individuals with their true digital identity are often used to establish GI, e.g., biometrics or PKI based on qualified certificates in the European context.

[4] Mehrdad Naderi et al., Towards a Framework for Federated Global Identity Management International Journal of Network Security, Vol.7, No.1, PP.88-99, July 2008.

2.5.3 Network Identity

According to the definition in "Strategic Implications of Network Identity",[5] a network identity (NI) is defined as the "context-sensitive identity, attributes, rights, and entitlements, all maintained within a policy-based trusted network framework. Managing Network Identity describes the software infrastructure and business processes for managing the life cycle and usage of an identity, including those attributes, rights, and entitlements". Up to now various network identity services (NIS) have been introduced. Popular examples are the Lightweight Directory Access Protocol (LDAP), Microsoft's Active Directory, Novell Directory Services (NDS), Kerberos, the concept of Public Key Infrastructure (PKI) as well as basic network protocols like the Domain Name System (DNS) or the Dynamic Host Configuration Protocol (DHCP).

In the last years, Network Identity Management (NIM) has advanced to further and more complex development stages [11]. The first stage is the "Identity Linking with mutual consent", i.e. identities still have unique and distinct profiles with SPs - e.g. online shops. However, identity profiles of different SPs are linked together, so that for example information/advertising material from third parties can be delivered. A second stage of NIM evolvement is the use of Identity Circles of Trust. In this stage IdPs come into play. SPs trust the identity provider, which provides the necessary identity attributes on behalf of the user. The last and most evolved stage is the use of federated NIS. In this stage we have a circle of trust between different IdPs. Trust relationships between SPs and IdPs are implicitly established by the circle of trust and its supporting initial contracts or community agreements. This is obviously the most flexible identity model, as users can connect with any IdP as long as it is a member of the IdP trust circle.

Currently the most used NI model on the Internet is that users have bilateral relationships with SPs (local network identity). Neither stage 1 is widely spread. However, especially in the last years we can observe the evolvement of stage 2, federated NI services, where users can login at web services using a hosted authenticator, e.g. Google Apps or Facebook. Other popular NIM approaches are Microsoft Live ID (formerly Microsoft Passport) or the Liberty Alliance project (now continued by the Kantara initiative).

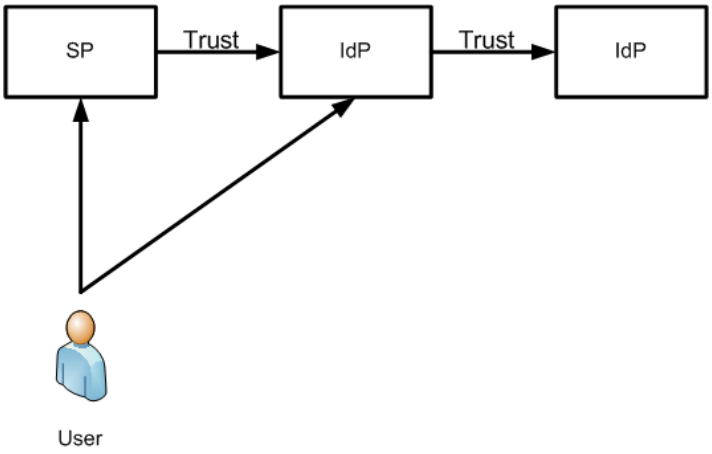
2.5.4 Federated Identity

Poetsch et al.[6] state that "Identity federation is based on a conceptual separation between service providers (SP) and identity providers (IdP) and concerns the arrangements that are made among several organisations and individuals, that let entities use the same sets of identification data, to get access (and authorisation) to the several different (otherwise autonomous) services offered by all the organisations associated with the system of federation". Thus, Federated Identity (FI) denotes the portability of identity information across multiple systems or organisations.

[5] "Strategic Implications of Network Identity", Sun Microsystems, http://www.sun.com/software/whitepapers/webservices/wp-identity.pdf.

[6] Stefanie Poetzsch, Martin Meints, Bart Priem, Ronald Leenes, Rani Husseiki "D3.12 – Federated Identity Management – what's in it for the citizen/customer". Deliverable D3.1 of Future of Identity in the Information Society – FIDIS, EC Contract No. 507512, June 2009.

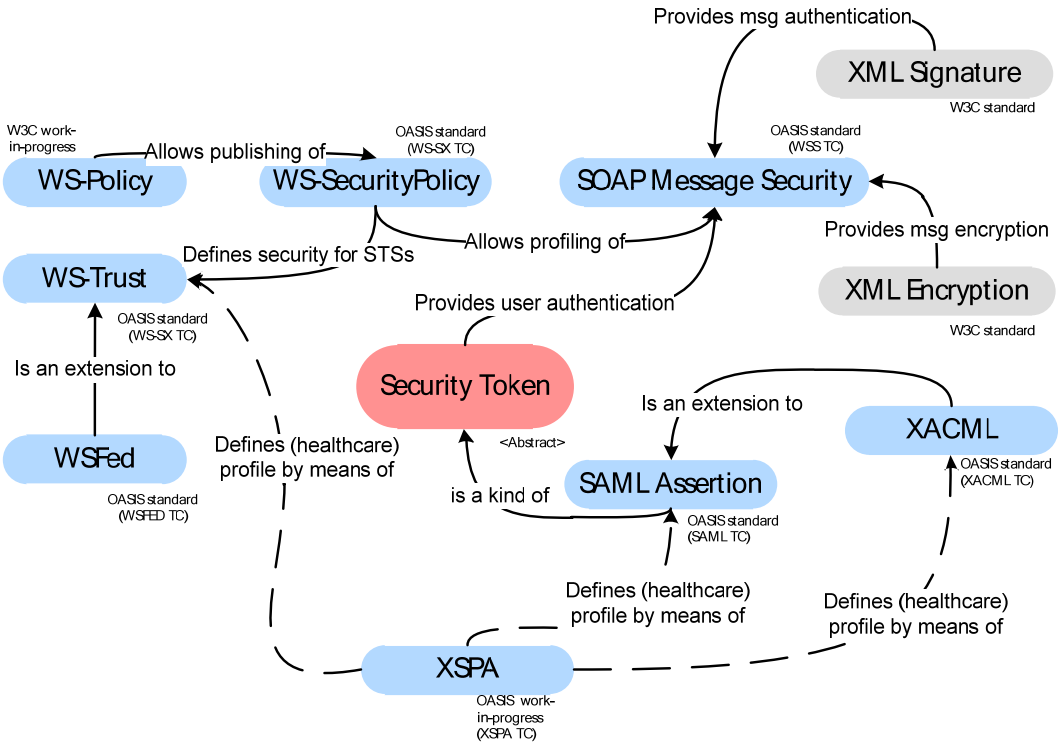
Figure 6: Federated Identity Model



In contrast to a local or centralized identity, the identity information is not stored within the same domain or network. As illustrated in Figure 6, the storage of identity information is located at so-called identity providers (IdP), which own or host the identity data. The goal of FI is the seamless edata access by users outside their own domain. True identity federation enables Single-Sign-On (SSO) across organisational boundaries. The concept of SSO using Federated Identity Management (FIM) allows users to automatically login in multiple SP applications by just authenticating once at the IdP. The exchange of identity information between IdPs and SPs is carried out using IF standards and technologies.

The most popular IF standard is the OASIS Security Assertion Markup Language (SAML). The Liberty Alliance, an industry group of more than 150 members, provides an FIM, which also uses SAML as a core standard.

Figure 7: SAML put in perspective



Other popular FIM frameworks are:

- WS-Federation;
- OpenID;
- Microsoft CardSpace (Information Cards);
- Higgins;
- Shibboleth;
- Central Authentication Service (CAS).

2.5.5 Brokered Identity

Like FI, a Brokered Identity (BI) is also part of the distributed identity model family. However, in contrast to FI, which uses a "web of trust" model, the BI model relies on trusted third parties (brokers) acting on behalf of the identity owner. In this case the identity broker is not to be confused with an IdP. It can rather be seen as a facilitator that is similar to an escrow service for managing online payments.

2.6 Centralised and Decentralised Identity Management

Since the management of identities is not a new a problem, various approaches for identity management systems exist. Some of those models are based on central storage of identification data; in contrast, others rely on decentralized or federated data repositories. Referring to a case study by Palfrey and Gasser[7], the next sub-sections briefly summarize the three main types of identity models. Basically, all three models follow the same approach having an identity provider, a service provider and a user who wishes to authenticate at service provider by the help of the identity provider. The distinctive criterion however is the ownership of the identification or user data.

2.6.1 User-Centric Model

In this model, the user always remains in control of her personal identity data. Personal data is only stored in the user's domain, such as his personal computer or a smart-card. Thus the identity provider is close to the user. If requested, identification data is transferred to a service provider via an identity provider only if the user explicitly gives his consent to do so. In fact, the user always carries the responsibility for releasing his personal data.

[7] Palfrey J., Gasser U.: Digital Identity Interoperability and eInnovation, Case Study, November 2007, Berkman Publication Series.

2.6.2 Centrally-Controlled Model

The centrally-controlled model is currently the dominant approach for identification and authentication on the Internet. At most websites – before being able to use services – users are asked to provide personal information for registration. This personal information is typically stored in central repositories in the identity provider's domain. Service Provider and Identity Provider are often located in the same domain.

2.6.3 Federated Model

In contrast to the centralized approach, the identity data is distributed across various domains and providers that have a trust relationship amongst each other. Each identity provider only stores portions of the identity information that exists for the user within the federation. Due to the trust relationship and special linkage of data repositories, identity information can easily be shared and distributed amongst the identity providers. In this model, the user's identity is referred as federated identity (see section 2.5.4).

2.7 Identity 2.0

The term Identity 2.0 is derived from the wording Web 2.0 that depicts the change from an informational World Wide Web to a more collaborative World Wide Web. Identity 2.0 has been introduced due to the heterogeneity of digital identities in the Web. Currently, users need to register or authenticate at each website if they want to consume certain services. If username/password authentication mechanisms are used, they also need to come up with a username for each website and they need to remember the passwords for them. However, mapping to the context of identity, each service provider or identity provider stores parts of a user's identity or even the same identity information (e.g. the user's mail address) is multiply stored in different repositories. Single Sign-On is one concept to overcome this issue of multiple identification and authentication (see section). The federation of identity data is another approach for the linkage of identity information.

Identity 2.0 also describes a possible solution against this identity data heterogeneity but focuses more on user-centric technologies. The idea is not to have multiple usernames and distributed identities but only one identity - under user control - to be used in various online situations.

2.8 Conclusion

This chapter has introduced the primary principles, patterns, and processes characterizing the technical aspects of managing digital identities. This serves as a basis for the use-cases presented below in Chapter 5, as well as the INDI architecture developed in D2.1 – "Logical Outline of the INDI Service Framework", and the privacy framework in D4.1.

Extending from this chapter, there is a basis for advancing towards working out the basis for assessing which technology offers opportunities for exploitation in support of the INDI vision. In the following, we consider the implications for the development of a privacy-friendly ecosystem for identity management, including concrete security requirements and safeguards

3 Privacy, Trust and Security

In computer science today, there is a plethora of definitions of privacy used in developing privacy solutions. These rely on different sets of assumptions and have different objectives.

In one school of thought, privacy can be defined as control over personal data, relying on the trust in third parties as long as they enforce certain policies, fulfil their data protection obligations, and introduce transparency mechanisms. This model of privacy, which we call "privacy as control", depends on the successful enforcement of data protection legislation e.g., EU Data Protection Directive, Fair Information Practices. Such legislation has the objective of increasing the accountability of organisations by requiring the data collectors and processors to be transparent about their data collection and processing activities, and by installing procedural mechanisms to exercise oversight over data collectors and processors.

In another school of thought, privacy can be defined as guaranteeing confidentiality of data, which requires the minimization of the data collected in systems as well as the minimization of individuals' need to trust third parties. These approaches are more preventative, rather than reactive and remedial. Typical implementations of this school of thought either conceal the identity of the individual user to the service provider, if identification is needed, minimize the amount of data revealed, or in even other cases, provide both concealment of identity and minimization of data revealed.

Ideally, the privacy as control mechanisms should also minimize data collection and allow individuals to choose from a palette of privacy properties, e.g., anonymity, pseudonymity, unlinkability of identities. In practice, however, most organisations will rely on strong trust assumptions and push access control and transparency to the foreground while avoiding data minimization.

Similar trends can be found in the different types of identity management solutions. In solutions proposed early on, e.g., Kerberos, when a user wants to authenticate to a relying party, he requests a temporary key from the identity provider. The identity provider sends the user two tokens, one temporary key for the user that may also contain some attributes, and another one for the relying party. This architecture has the undesirable consequence that each time the user makes a request for a service that requires authentication, this is known to the identity provider, leading to numerous privacy risks.

Later solutions suggest addressing this problem by introducing a number of identity service providers. In this "federated" model, a user that wants to authenticate with a relying party first authenticates with an identity provider of his choice. Upon successful authentication, the identity provider generates a signed token that certifies the identity or other attributes of the user, and a key (pair) that the user uses to authenticate himself with the relying party. In this architecture, any relying party with an authenticated copy of the identity provider's public key can authenticate the users. Although the introduction of multiple identity providers addresses some of the privacy risks that occur in the case of a single identity provider, this architecture does not solve the problem that each service provider will have to manage and secure a large database, will be able to profile all of its users, and colluding identity service providers will be able to enhance these profiles. Given that economic incentives have so far overridden privacy concerns, and sharing of datasets between data collectors and third parties has only surged in the last decade, collusion and incompliance remain a substantial risk in this proposed architecture.

There are a couple of solutions to the privacy risks inherent to the federated identity provider architecture. First, users may cache tokens and hide the number of times they use a service, although they will not be able to hide the fact that they use the service. Further, they may carefully

generate identities at multiple identity providers so that even if these collude, they will not be able to link these different identities.

A solution that further mitigates the existing privacy risks depends on the use of advanced cryptographic systems such as group signatures, anonymous credentials and zero knowledge proofs. In this model, relying parties, users and identity providers need public-private key pairs. When a user registers with an identity provider, he receives a credential from the identity provider containing her attributes. Each time the user wants to interact with a service provider, he will use the credential to prove a logical statement about the attributes in it. He will not hand over the credential, but only prove the absolutely necessary information to the service provider and nothing else, e.g., he will be able to prove that he is over 18 and has residency in a given location. Assuming that anonymity is not compromised at other layers (e.g. the network layer), the service provider learns nothing other than the desired statement. In this architecture, the identity providers only know about the credentials they provide to the user, while the service providers only learn the minimally necessary information, addressing each of the privacy risks raised with the federated and monolithic identity provider models.

Further solutions are being proposed. A typical example is OpenID, which allows users to assert self-certified claims. Another interesting example is OAuth, which is a user controlled mechanism to delegate access to resources, e.g., allowing a social network service provider to access the address book at the webmail service provider. These solutions are interesting in that they allow for the identity matter to be separated from other content, while not reproducing the privacy risks associated with identity providers. Nevertheless, these models also have their shortcomings for those assertions that are better asserted by authoritative sources.

While the monolithic identity model and the federated model are typical examples of privacy as control, the architecture with anonymous credentials is a mélange of both the privacy as control as well as the privacy as confidentiality model. We have yet to see a broad deployment of these technologies. The experience is likely to lead changes in some of the assumptions with respect to how to best enable identity management and protect privacy of individuals concurrently.

In the rest of this chapter we discuss some of the identity threats inherent to the design of the Internet and to current proposals for and implementations of identity management systems. We then provide a short overview of the privacy and security requirements that have been suggested by computer scientists to mitigate some of these threats. In the last section, we provide an overview of the privacy, security and trust mechanisms proposed by solutions in the federated identity management model.

3.1 Identity Threats

The following sections provide a brief overview on a selection of common threats regarding digital identities.

3.1.1 Security of Identities and Surveillance

Many digital identities are central to the users' everyday activities. Hence, physical loss of a digital identity (loss of availability) is one of the gravest threats to the identity model. Just as strong as the loss of availability is the loss of the integrity of the digital identity, e.g., the identity may seem to have performed actions that the user never took, or vice versa. Further, if there is a confidentiality breach with respect to the identity data, then unauthorized others have accessed the identity and may process or distribute the data of the given user.

From the perspective of law enforcement, governments and businesses, these security properties are also important. Breaches to the availability and confidentiality of a user's identity data may be

detrimental to the services offered by the organisation, to activities of the organisation itself, as well as third parties who rely on the data. Integrity of the digital identity, e.g., the uniqueness of a digital identity, is also of considerable importance to organisations and their transactions.

Given that identity data is both of economic and organisational value, and given that identities will be managed by these organisations with economic and organisational interest, the interest in keeping the security and uniqueness of the ids is likely to be better protected by organisations and government than the sometimes conflicting security and privacy requirements of the users. Given their power advantage and omnipresence, this is likely to lead to issues with respect to unnecessary collection of identity data, as well as limited access to services for those users that do not provide their data. Such conflicts have to be addressed such that priority is given to user privacy and security needs. Organisations are unlikely to have incentives to provide users with such options, meaning that policy and regulation may be necessary to guarantee that identity management systems do not become by default intrusive surveillance systems that coerce users into providing their personal data disproportionally and excessively for the purpose of data collection.

By surveillance we are referring not only to the surveillance of individuals but also of communities and populations. By having access to the profiles and behaviour of millions of users, identity providers are likely to practice statistical analysis of their user base. Such analysis of populations of users in order to identify different categories of users, some of which will be privileged while others are discriminated, is what is known as the social sorting and discrimination problem resulting from surveillance. Current identity architectures try to solve this surveillance problem by allowing users to select from multiple identity providers or using anonymous credentials. The obfuscation resulting from the use of numerous identity providers does not protect the users from social sorting, while anonymous credentials may provide some protection by minimizing the collection of population data. Regardless of the amount of data that substantiates statistical surveillance and profiling, methods and regulation to limit statistical surveillance, to mitigate resulting discrimination, as well as making such systems transparent, accessible and refutable must be put in place. The INDI model may play a leading role in making such efforts possible, if it is truly driven by user interests.

3.1.2 Identities over Time

A rather problematic development in the current digital service delivery is the fact that technical means are implemented much faster than adequate regulatory protection and awareness means. Hence, the following issues that can be categorized as social threats are relevant in our current ecosystem of information systems and regulations.

The imagined time span for use of information disclosed to digital information systems may in reality not hold. For example, information revealed to a service provider may potentially be – in contrast to traditional paper record keeping – available for an indefinite period of time. While a service consumer formerly may have relied on the fact that any piece of information could be permanently destroyed, modern systems feature a worrying trend towards keeping everything forever.

This fact alone raises privacy and data protection questions and may lead to formerly unknown social threats. In particular social networks, such as Facebook, Myspace, and similar organisations, gain full control over all user-generated content at the point of disclosure and may trace every action any service consumer performs at the respective social platform and its associated network. Traditional data protection means and legislation are only partially applicable to those new services due to several factors that are inadequately reflected as of yet.

First of all, the participation in any social network is fully voluntary as well as the provision of user-generated content. However, the content, as soon as published on social platforms, may be

processed freely by the platform. That may result in an unintended, collateral data disclosure, even or in particular for "private" content as illustrated by this exemplary ruling: "*A plaintiff must give a defendant access to private postings from two social networking sites that could contradict claims she made in a personal injury action, a Suffolk County, N.Y., judge has ruled*". The same ruling also revealed that all user-generated content on social platforms is actually systematically stored and may be objected to unwanted disclosure by requesting "*current and historical records/information*".

The opposite is also possible: information revealed to a service provider may be lost once those services are cancelled, e.g., GeoCities with 38 million user-built pages was taken offline after 15 years. This loss of data becomes an important threat for information that was not made public, e.g., public profiles in social networks, but which consisted of information shared between the user and the service provider, e.g., personal or historical information important to the individual user.

While the never-forgetting and amnesiac internet are both relevant but extreme cases, the general problem of the validity of an identity over time with respect the expectations of users and service providers is likely to prove a difficult theme. Both the revocation of identities as well as their prolongation and updates may not be well synchronized across the Internet, causing severe damages to individuals but also to the reputation and reliability of identity architectures. This problem becomes especially acute when it comes to biometric information, which is not easily revocable or renewable.

3.1.3 Incapacitated Data Controllers

User-generated content, such as posts, pictures, and personal preferences, published by the user on social platforms is assumed to be still exclusively owned by the user in the popular opinion.

According to the New York Law Journal article mentioned in the last section, the judge ruled that the user "*consented to the fact that her personal information would be shared with others, notwithstanding her privacy settings. Indeed, that is the very nature and purpose of these social networking sites, or they would cease to exist*".

Furthermore, the data controller's (user's) ability to permanently delete content may be heavily impaired on many social platforms. While it is still possible to hide the display of current or historical information disclosed on a platform, the actual data usually remains unaffected and permanently stored within the respective platform: "*Even after you remove information from your profile or delete your account, copies of that information may remain viewable elsewhere to the extent it has been shared with others, it was otherwise distributed pursuant to your privacy setting, or it was copied or stored by other users. However, your name will no longer be associated with that information on Facebook. (For example, if you post something to another user's profile and then you delete your account, that post may remain, but be attributed to an "Anonymous Facebook User.") Additionally, we may retain certain information to prevent identity theft and other misconduct even if deletion has been requested.*[8]"

As a result, potentially there is a serious degradation of the rights to data assigned to a data subject.

[8] Facebook, Privacy Policy, http://facebook, com/policy.php, last accessed: December 10th, 2010.

3.1.4 Over-Sharing of Personal Information

Today's digital services offer a vast variety of useful applications. However, most services require a registration and the subsequent assignment of a digital identity before any service may be consumed. While each identity considered separately usually is not expected to pose any threat, a simple combination of many may quickly turn a service consumer totally transparent, rendering formerly assumed user rights, such as anonymity, privacy, and confidentiality, ineffective.

Potential threats to any of the user's rights may arise from formerly unsuspicious "every-day" technology, such as navigation systems, mobile phones, and RFID-enabled objects (ranging from clothing to government-issued I.D. cards). Navigation systems usually feature a roadside monitoring service, for instance a traffic congestion advisory system as a value-added subscription. In order to deliver the latter service, the system obviously needs to constantly disclose the car's current position and projected destination in order to acquire the correct traffic status information.

As already briefly mentioned in the previous section, in particular social platforms experience a huge increase in rather questionable re-use of the information published on them. One example is a survey performed by the American Academy of Matrimonial Lawyers, which revealed that "*An overwhelming 81% of the nation's top divorce attorneys say they have seen an increase in the number of cases using social networking evidence during the past five years*"[9].

According to this study, 81% of AAML's members have seen a significant increase in using evidence derived from social platforms and AAML's members observe a potential over-sharing of personal information in digital services, which may cause significant disadvantages in real life: "*As everyone continues to share more and more aspects of their lives on social networking sites, they leave themselves open to much greater examinations of both their public and private lives in these sensitive situations*".

3.1.5 Impairment of Rights by Digital Services

Digital service provisioning in conjunction with today's inadequate protection means may also impact certain actors that are explicitly protected by law, such as whistle-blowers, whereas whistleblowers shall be generally protected from retaliation and shall remain anonymous.

3.2 Over-use of Identity Data

3.2.1 Claimed/Requested Attributes

In particular digital identities issued by a multi-purpose IdP usually feature a tendency of assigning a maximum of attributes in order to support many use cases and scenarios straight "out of the box". Since the user's actions are usually unknown prior to their occurrence, the assigned attribute set might be significantly bigger than actually required. Generally it is considered good practice to initialize the digital identity with a very slim set of assigned attributes and to provide additional attributes on an "on-demand" basis. A typical scenario is an individual serving in a dual

[9] American Academy of Matrimonial Lawyers, **Big Surge in Social Networking Evidence Says Survey of Nation's Top Divorce Lawyers**, http://www.aaml.org/about-the-academy/press/press-releases/e-discovery/big-surge-social-networking-evidence-says-survey-, last accessed: December, 3rd, 2010.

role in an organisation, such as a medical doctor who is also assigned with supervisory functions vis-à-vis other employees. Even for performing his usual tasks as a medical doctor, the system assigns him the role of supervision, which may lead to a much bigger data disclosure than actually required by his regular duties.

The opposite situation may occur when a service consumer requests a maximum set of attributes regardless of whether all attributes are actually required in order to fulfil the requested action.

3.2.2 Documentation Requirements

Almost all digital service providers need to produce sufficient evidence regarding their own service operation and about the service consumers' actions. This requirement is usually directly derived from the regulatory framework which as applicable for the concrete service provider.

For instance, an e-business service provider is required to keep financial records for a certain period of time, including the information about who acquired what at which time and in which environment. A medical doctor on the other hand is required to keep all relevant documentation regarding his treatments, including all medically relevant information about all treated patients. Naturally this requirement also affects digitally communicated indications, treatment plans, and other types of medical information. Even more, not only the medical documents need to be properly conserved but also the concrete communication environment such as access dates, treatment relationship, activated roles, and consumed side-services.

3.3 Security and Privacy Requirements

Privacy and security engineering research proposes a number of requirements that need to be fulfilled to address and mitigate some of the identity threats identified above. In the following section, we provide a brief overview of these requirements.

3.3.1 Security Requirements

In information security, the following core concepts are traditionally used to describe the primary protection demands.

- **Confidentiality** describes the property of a system to avoid the unauthorised disclosure of information.
- **Integrity** describes the property of a system to ensure that information may not be manipulated without detection.
- **Authenticity** describes the property of a given system to ensure that all entities, provisions, and assurances that are required and processed during any data processing are genuine.
- **Availability** describes the capability of a system to ensure that the required information is available whenever required and subsequently that it is able to perform its assigned tasks within an acceptable time frame.
- The term **Non-Repudiation** addresses the capacity of a given system or entity to ensure that the actual execution of a given event cannot not be successfully disputed in retrospect.

3.3.2 Privacy Requirements

The following privacy requirements can be listed as refinements of the security requirements. However, in multi-stakeholder systems while security requirements represent requirements that may be common to all stakeholders of the system, e.g., all stakeholders would want the system to be secure, privacy requirements are specifically central to fulfilling user needs, and hence user-centric. The following are the most prominent types of privacy requirements, for which privacy technologies are also in development.

- **Anonymity**: is the condition of not being identifiable within a set of subjects. The anonymity set for a given action is the set of all subjects who might have triggered the action.
- **Pseudonymity**: is an identifier used in place of the "real" identities, e.g., name, unique id number, of a given user. Pseudonymous identifiers can be made conditional and accountable using cryptographic building blocks. Simple forms of pseudonymity, where the user replaces his or her own id or does not provide an id, but other traces can be used as a pseudonymous handle, provide the user with little protection that relies on the obscurity in the system.
- **Unlinkability**: is the condition in which a third party cannot determine whether two actions or two data items belong to a single user. Unlinkability is central to another privacy related concept called the separation of identities.
- **Separation of Identities**: the condition of guaranteeing that separate partial identities of a given user are unlinkable.
- **Separation of Audiences**: the condition in which a user can control the audience of the information s/he reveals. The flexibility of the access control models determine the type of separation of audiences that can be practiced by the user.

3.4 Identity Mechanisms for Security, Privacy, and Trust

3.4.1 Roles and Delegation

Modern IT systems feature multiple users, resources, actions, and contexts. The potential product of every property may lead to an unmanageable amount of access control definitions and decisions. One consolidated means to reduce the complexity of a system is to derive common access patterns/demands and to subsequently assign every system user to one or more of the categories: the roles. As already described in the previous chapter, it is common good practice to further decouple the security and safeguard means. Therefore, a role traditionally consists of two parts, the role definition and an assignment.

The role definition consists of the role's name and a distinctive set of rights to be associated with that particular role and bound to a specific object. The role assignment represents the relationship of the role definition to one of the system's subjects. The separation of those two aspects avoids the direct assignment of rights to a specific subject's identity.

The concrete execution of access rights is therefore not immediately bound to the user but implicitly acquired through its current role at the specific time any action is requested to be performed. The roles and their associated permissions may be defined hierarchically and rules may be constructed which define limitations (constraints) for the role assignment and permissions-granting.

3.4.2 Rights Delegation and Mandating

Permissions in digital systems are traditionally assigned following the "least privilege" principle. While that principle fits perfectly in theory, it is only partially applicable to real-world scenarios because of many inconsistencies between the technical means and the organisation of work in practice. In many real business services, the human actors and security safeguards allow for a technical representation of an informal, permanent or temporary authorisation for a specific action granted by the rights bearer upon particular request.

In the real world this might be a doctor who temporarily authorises a nurse to check a formerly-inaccessible patient summary for the next follow-up appointment, or a head of medicine empowering another healthcare professional to have access to all formerly-inaccessible medical documents of an organisation for patient safety reasons.

In more generalised technical terms, this may translate to: "making it possible to express permissions about the right to issue policies and to verify issued policies against these permissions".

One subset of the delegation principle is a mandate. While a delegation may optionally feature a distinct assigning/assignee relationship between the original rights bearer and the recipient, the mandating is built upon a direct and explicit assignment of the respective rights and functions. Additionally, a right exercising model in which the recipient is instantiating rights to perform an action under the responsibility and rights composition of the original rights bearer may be observed. In a mandate scenario:

- the rights recipient organisationally performs all qualified actions under the capacity and responsibility of the original rights holder with no additional assignment of any rights (acting on-behalf-of);
- in a mandate scenario, the rights recipient does not possess the capacity to extend the assigned set of rights (e. g. shall not issue new policies but merely instantiate already existing polices of the original rights bearer).

An example of mandating may be sketched as follows: In order to enable a Spanish nurse to perform her daily duties in full compliance with the local organisation of work and regulatory framework treating a patient whose data is controlled by an entity from another regulatory domain. In order to be eligible to gain access to the foreign data, she may be organisationally and technically acting on behalf of a medical doctor in order to comply with the foreign access policies.

3.4.3 Policies

Policies in general are considered to define the rules and borders within a given system, thus to govern the behaviour of a given system. Initially defined as rather abstract rule sets on paper, policies are then usually transposed into a technical representation – the policy language – in order to automatically enforce the systems behaviour to be in full compliance to the policies provisions.

A typical policy consists of:

- target identifiers, that denote which (kinds of) entities (resources, subjects, etc.) are regulated by the policy;
- conditions, that define rules for applicable targets:
- rules and constraints, for deciding on whether a certain permission is granted or not;

- obligations, that have to be processed in order to fulfil assurances that are associated with certain permissions.

Due to the rather generic and very flexible composition of a policy[10], policies may be utilised in order to express nearly any constraint related to the IT-based processing and decision process. Given the manifold sources (e. g. legal, regulations, privacy consents, ToCs) for these constraints, one single policy for expressing each constraint may be very complex and too specific. Therefore, it is highly advisable to firstly identify the different policy concerns and define separate policies for the individual concerns that can then be stacked and combined accordingly whenever needed.

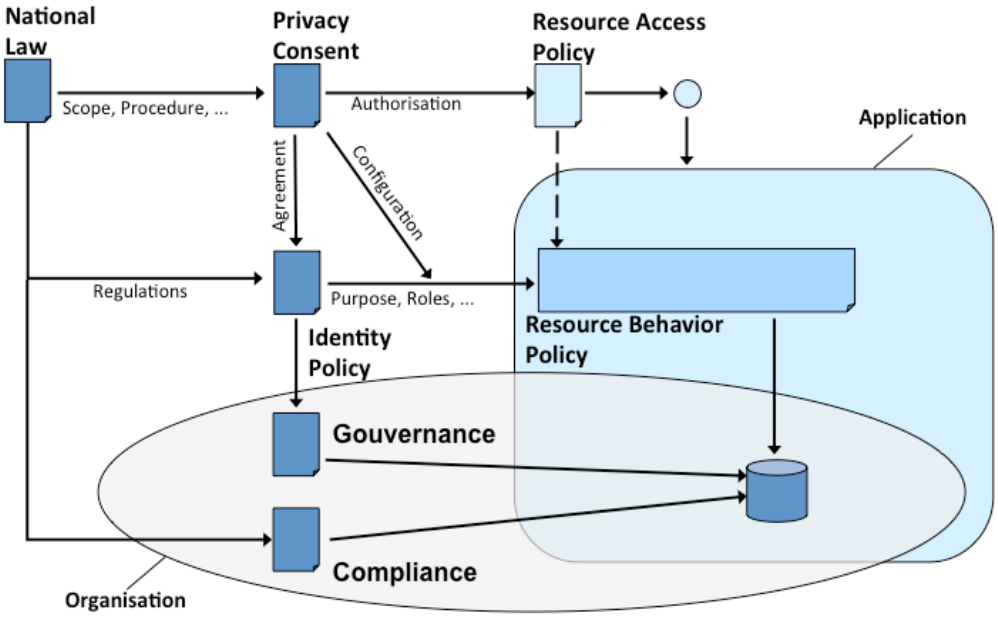
With respect to identity management, the following major concerns need to be considered:

- legitimate uses and acceptable purposes of operating digital identities, based on the concrete legal and regulatory environment;
- privacy consent: constrains the possible uses the data subject puts on the use of his data;
- application semantics: constraints that can be derived from the purpose of use of a certain application that mediates (or even initiates) an access to attempt a protected resource;
- compliance: resource security rules for protecting identity data within an organisation from illicit disclosure and use;
- the extent to which complementary and environmental policies are supported.

Figure 8 illustrates the separation and relationships of policy concerns. It also shows how these concerns relate to each other and in what ways they correspond to policies that have to be enforced in conjunction with the processing of identity data.

[10] How much of this flexible is really available, depends on the expressiveness of the policy language used.

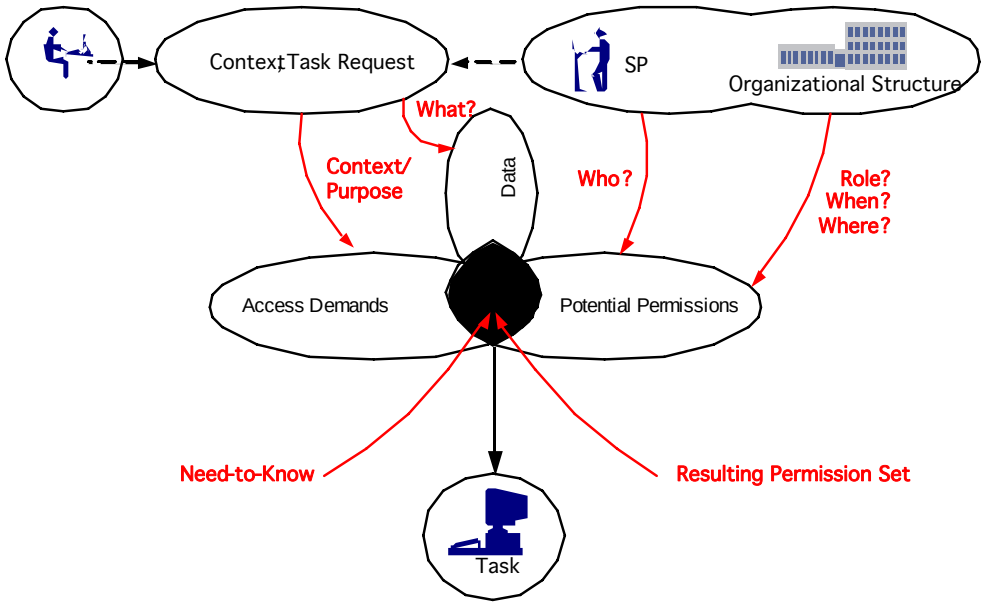
Figure 8: Separation and relationships of policy concerns



Additional to the separation of concerns principle, other best practices exist in reference to policies. The two most important principles are "need-to-know" and "least privilege". Both are designed to explicitly sanction the service provider's disclosure scope and liberty of action.

While the "need-to-know" principle primarily regulates the actual amount of information disclosed to the service provider, the "least privilege" principle enforces the granting of only the lowest composition of permissions that is required to execute the assigned and authorised task.

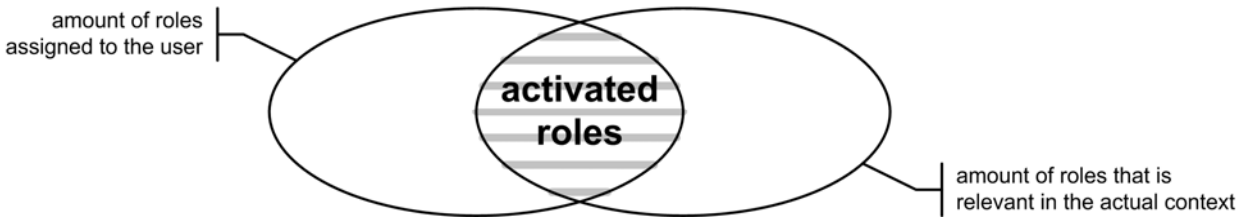
Figure 9: Need-to-know principle



The "least privilege" principle primarily regulates the number of roles activated and permissions assigned to each service provider. To each individual actor of a service provider several (one or many) roles may be assigned, depending on the current work context and the organisation of

work. In order to follow the design principle of least privilege, the IdM and access control system must ensure that for each actor only and exclusively those roles are currently activated which directly correspond to this person's momentary activities.

Figure 10: Restrictive role activation



The number of applicable roles is determined by calculating the intersection of the user's theoretically assignable roles (all roles administrated for him in the subject domain) and the roles required to act in the current context. The activation (identification) of the current context is usually an implicit side effect caused by actions such as switching applications, assigned tasks or context.

3.4.3.1 Identity Policies

In contrast to the other, more technical policies described in the following sections, identity policies usually regulate the behaviour and rules of conduct within a system or federation of systems. The identity policy furthermore directly reflects the chosen architecture (legally and technically) of the system as well as the conceptual foundation for all subsequent identity management and enforcement strategies.

Due to its umbrella characteristics the identity policy is of outstanding importance in any IdM system by integrating and sanctioning most of the other functions such as technical policy provision, processing, and enforcement: In many systems of considerable size, the identity policy is the first element in a control- and data chain that has to be traversed with each service request. This chain is usually called policy stack and illustrates the strictly arranged, consecutive orchestration of services and safeguards within a system.

In summary, the identity policy:

- defines the acceptable behaviour of the system or federated system;
- incorporates and enforces the constant application of the legal and regulatory framework in which the system is operating;
- reflects the architecture and coordinates the orchestration of the subsequent IdM and security services;
- communicates the capabilities, constraints, and assurances of the affected system when operated in a federated environment;
- calls for the specific procedures for the individual tasks to be performed.

Implementation

Defining and implementing adequate identity policies is considered to be a rather challenging task that is traditionally circumvented by strictly concentrating on a certain object (resource) or task

(action) instead of the individual (subject). This commonly results in digital identities and attached rule sets that are designed to be exclusively applicable in one domain, such as traditional identity cards for the identification of citizens by their government.

In reality, however, the broad availability, already existing infrastructure, and comparable cost-effective dissemination of such identities (usually paid for by the individual) create a strong demand for a cross-domain re-use of those formerly restricted domain applications. Such behaviour often results in inadequate and inconsistent identity policy application, such as the unintended disclosure of personal information or the abuse of an identification means (such as an identity card) as authorisation means towards a third party.

3.4.3.2 Identity Assurance Policy

In the traditional, paper- and "thing"-based world, the determination of a subject's identity is usually facilitated by government-issued token, such as an identity card, passport, or a driver's license. These tokens are considered to feature a very high assurance level, which simply means that a relying party may trust the identity of a claimant with a very high degree of certainty.

In the digital world, identity tokens with a degree of assurance concerning the factual identity of the claimant do exist but may not be used for all purposes due to their current design, such as cost-effectiveness, proportionality of the means, and privacy concerns. However, a large number of other digital identities with various degrees of assurance may also exist that are better suited to be operated in a particular scenario.

In summary, different business scenarios require different means, authenticity levels, and stability. While a rather basic identity determination may be sufficient for many services, some require a high degree of confidence. In order to formalize those specific requirements and to act as a facilitator for automatic processing and enforcement, the required or asserted degree of certainty may be defined in a special policy.

3.4.3.3 Privacy Policy

In order to lawfully collect, store, process, and communicate information about an individual, a concrete and prior authorisation is required for those operations. This authorisation is also referred to as privacy consent. It is the result of an individual's independent and informed decision and specifically defines:

- which of his data may be shared;
- through which applications and for what purposes;
- to what extent (partly, context-dependent, all);
- with whom (identities or organisations) the data may be shared;
- for how long?

Finding a suitable technical representation of the privacy policy is considered to be quite challenging due to the potentially high complexity of an adequate reflection of the concrete individuals wishes.

The specific rules and regulations, which are encoded in this policy, directly reflect all explicit and implicit authorisations that may result from the individual's decisions. Simply spoken; a privacy policy expresses the individual's choice on whom he trusts and what he is willing to share.

While application semantics and compliance-related policies mainly control how authorized users process protected resources, a privacy policy usually focuses on determining who is - at all - allowed to access the resource.

3.4.3.4 Resource Behaviour Policy

A resource is always managed within the context of an organisation that is liable for the lawful processing of its client's data. It is the responsibility of IT-compliance functions to define roles, permissions, and obligations for internal and external data communication: The respective resource behaviour policies are enforced whenever an access to an internal resource is requested.

3.4.3.5 Resource Access Policy

Based on the privacy policy and its translation and the current context, the authorisation of certain individuals, organisations, and/or rules to use the application with respect to the agreed semantics, are defined in the resource access policy. It controls who – if anyone – is able to access a protected resource within the context of a certain application.

3.4.3.6 Application of Policy Frameworks

The distinct separation of the policy concerns and the ability to build flexible policy processing stacks – policy frameworks – that may be combined to foster benefits for the user, such as:

- increasing information security aspects of a service, in particular within loosely coupled federations;
- fostering public-private partnerships;
- exploiting re-use capabilities and improving cost-effectiveness;
- getting the stakeholders involved and move them into a position in which they may formulate and enforce their wishes.

3.4.4 Trusted Third Parties

In the traditional organisation of work and business transactions, it is a common practice to interact with certain intermediaries to safeguard service deliveries, such as a notary service for authenticating the affected parties and a subsequent vouching for their authorisation by witnessing the provision of the wet-ink signatures on a piece of paper.

In certain digital scenarios, such as pseudonymization, evidence-keeping or escrow services, and the secure provision of cryptographic material, a service provider and a consumer may also agree on decoupling their direct relationship by interposing an entity that is trusted by both, service provider and consumer. Such an entity is traditionally called a Trusted Third Party (TTP) and frequently operated to certain trust extents in a digital context.

TTPs exist in many flavours in the digital world, ranging from rather simple anonymizing services to advanced payment providers to highly sophisticated pseudonymization services for enabling medical research on real patient data with re-identification means in case a cure has been found.

However, despite their apparent benefits for digital services of many kinds, TTPs are also faced with severe concerns primarily by data protection and data safety entities. By design, a TTP is always threatened by representing a single point of failure and is therefore considered to be a

pristine vector for potential attacks: When the TTP availability is impaired; the services relying on the TTP's operation may not be delivered. Furthermore, while interposing a TTP may greatly reduce the processing of sensitive information within the service provider and consumer realm, the TTP itself needs to document all transactions and communication as well. A successful attack on the TTP confidentiality may therefore disclose the information about both parties at once, service provider and consumer. As a result, TTPs must be thoroughly protected.

Another significant concern regarding TTPs in the digital domain is the inability to technically assess further constraints on transactions, such as the actual willingness, regulatory validity, and absence of pressure. A patient privacy consent may serve as an illustrative example: In digital health services, a patient's consent is: "*any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed*"[11]. The patient usually agrees by signing digitally using his own cryptographic key material protected by an unshared secret (holder-of-key). The TTP, most likely a public key infrastructure with digital signature functionality in this case, checks the correctness and validity of the material provided and subsequently grants the request. However, the patient's authorisation is merely implicit, since the CA is unable to assume the validity and value of several mandatory properties, such as "freely given" and "informed".

Technical means may not adequately evaluate the patient's state of mind and his ability to understand, and therefore a traditional holder-of-key authorisation may only possess limited significance expressing that all patient rights have been fully respected. This may lead to real-world threads in situations in which the service consumer is the data controller and is forced to disclose information against his will and in violation of the data protection regulations.

In this Chapter, we provided an overview of the different types of architectures that may come into question when defining the GINI infrastructure. We discussed some of the threats to INDI identities and formal privacy properties and privacy by design principles that are relevant in the INDI space. The analysis of privacy threats and requirements is the topic of Deliverable 4.1. Further, we discussed access control and trust models that are relevant to the GINI infrastructure. We will further elaborate on these models in the deliverables of WP2 and WP4. These deliverables will include a gap analysis of the future research and implementation needs that are necessary for the INDI environment described in Chapter 5 to become a success for businesses while becoming an exemplary alternative to current personal data collection and processing practices.

[11] 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.

4 Common Patterns and Building Blocks

This chapter exploits the concepts provided by chapters 2 and 3 in order to describe stable standards, best practices, and innovative approaches in the field of digital identities.

In particular the emphasis on stable, international standards is of crucial importance for the provision of interoperable and sustainable identity management solutions. As a result, this chapter illustrates how state-of-the-art identity management services may be aligned to standards and how those may be put in practice successfully by using the respective standards and best practices

The depicted standards and best practices are also accompanied by real-world examples of their respective current implementation and application by recent projects, current initiatives, and services. Additionally to this information, common building blocks and deployment patterns of identity management services are shown and assessed.

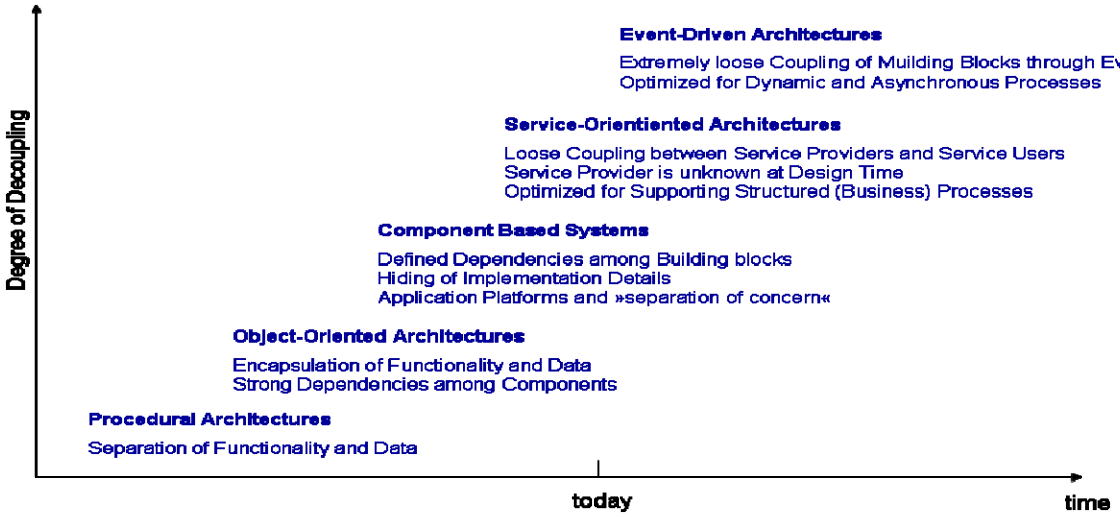
The provisions of this chapter are subsequently picked up as the foundation not only for this document and its contained use-cases, but also for the development of the privacy framework in D4.1 and D2.1 – "Logical Outline of the INDI Service Framework". All concepts and paradigms presented in this chapter are of vital importance for the more specific deliverables in those subsequent reports.

4.1 Decoupling Security Services

One fundamental principle of security related services, such as authentication and/or authorisation, is to adequately safeguard the privacy, integrity, reliability, and legal stability aspects of the subsequent business services, as well as their particular governance rules. In order to seamlessly and robustly address those responsibilities, the respective security services technically need to be either tightly linked or, if possible, closely integrated into the business IT services architecture that is to be protected. Obviously, the collaboration between the security and the business services work is proportionally beneficial the more both service paradigms match with each other.

As architectural paradigms and their particular sophistication evolve over time – with service-orientated architectures being considered to be the state-of-the-art at the moment – one individual aspect is shared between all modern architectures: the distinctive separation of security, infrastructure, and business services. This aspect is currently further refined by decoupling even the individual building blocks of the above named services in order to distinguish and independently operate sub-services and/or -components.

Figure 11: Decoupling as a core architectural paradigm



In the current best practice this leads to IT architectures in which security tasks are commonly decoupled to a degree that enables that task to be executed completely independently of the business services, with respect to service deployment, data consumption, execution place, and partially the execution time. For instance, security tasks such as authentication and authorisation may be totally externalised with only the tasks results – such as an authorisation decision – being made available to the requesting business service. Subsequently, the functionality of the decoupled tasks are masked behind distinctive and standardised service interfaces in order to enable a common consumption pattern and a flexible service orchestration.

The decoupling and service-encapsulation may greatly facilitate:

- privacy and confidentiality safeguards of business service since personal identification information is not automatically required to be seen and/or processed by the business services anymore;
- externalisation fosters security services interoperability by popular demand and may reduce operation costs;
- externalisation leads to heavy re-use of already existing identity data instead of the recreation of redundant data whenever a new service is to be consumed;
- seamless integration and service delivery, since by using decoupled security services means such as single sign-on may be applied, which avoid a duplicated execution of security tasks;
- reduction of complexity/redundancy: instead of using several identities each disclosing a set of personal information and being secured with separate means (PINs, TANs, passwords, etc.) only one service (not necessarily one physical service implementation) is comprehensively dealing with such tasks;
- re-location (stripping) of personal identity information from various services into one external service potentially controlled by the end-user;

- self-determination: end-users may decide on what personal information they want to provide instead of being forced to create new compulsory identities whenever they want to consume a service;
- traceability and transparency enhancements: what personal information is available and how is it being used;
- isolation of the individual security services increases the overall security by eliminating unintended collateral effects and correlations.

GINI follows this approach by defining security services for the management of identity lifecycle as business- and application-independent services. These identity services interact with business services through these services' security subsystem (see sections on "Single Sign-On" and "Circle of Trust" for examples).

Furthermore, GINI supports the alignment to technical and organisational international standards, the potential exploitation of cross-domain re-use of its services, and harmonised cross-border service delivery by separating its services, duties, concerns, and knowledge in its service implementation.

4.2 Standard Building Blocks

According to the decomposition of the security services as pictured in the former sections, the security services suitable for normal operation must at least provide the functionality to cover the following areas:

- management of digital identities and their respective life cycle;
- cross-domain trust-coordination, -establishment, and –modelling;
- management of identities, policies, and attributes;
- policy provision, decision, enforcement, and alignment;
- security token issuing and verification;
- semantic transformation and mediation.

Based on this compilation the following loosely coupled functional basic building blocks may be distinguished:

Table 2: Building Blocks of Digital Security Services

Attribute Service	An attribute service, that provides attributes which are a property, quality, or feature of a given subject that are the required information foundation for security-related decisions.
Authentication Service	An authentication service, that issues the assurance that a claim for a given property of a given subject at a given time is actually true.
Security Token Provider	A security token provider, that issues assertions which are small pieces of information (statements) whose correctness is assured and confirmed by electronic means, such as a digital signature.
Security Token Verifier	A security token verifier that validates safeguards of an assertion, for instance verifying the digital signature. A secure token verifier, however, does not verify the correctness of the included statements of an assertion.

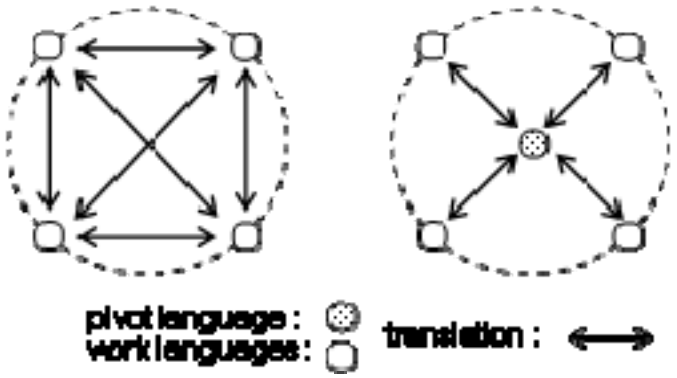
Trust Anchors	Trust anchors are actors of extraordinary legal stability or responsibility, such as national/international root certificate authorities, government primary identity providers, and / or supervision authorities.
Policy Enforcement Point	A PEP is an actor within a security service that intercepts all requests and all responses in order to ensure their full compliance with the current rule set and rule decisions applicable at the interceptions point in time.
Policy Information Point	A PIP is an actor within a security service that makes the applicable rule sets (policies) available for the subsequent policy processing and decision.
Semantic Services	Semantic services are generally considered to be mediators between two or more formerly incompatible (non-interoperable) actors by meaningfully connecting/relating the available and required information from two or more domains to each other.
Security Safeguards	Technical, regulatory, and organisation means to ensure compliance with the current governance rule set (see next section).
Identity Service Provider	An identity provider authenticates the claimed identity of a subject by verifying its provided credentials (e.g., a password, a signature, or a response to a challenge). After a successful authentication, the identity provider issues an identity assertion in its function as specialised secure token service.
Organisational Anchors	The foundation of all appropriate technical security and safeguarding means is an organisational setup that regulates the area of application, guarantees legal stability, and governs the fundamental trust between authoritative actors (trust anchors).

4.2.1 Semantic Services

Whenever identity services as sketched above are used in a decoupled way and are flexibly deployed, interoperability challenges arise. Electronic identities are highly interoperable and reusable since the pure electronic identity itself is carrying little to no meaning for the identity consumers. Meaningful content is traditionally added by the inclusion of mutual understandable attributes (enrichment).

However, attributes are highly domain- and application-specific. Specialisation however leads to interoperability issues since not every service provider is able or willing to comprehend all attributes, although the attributes in question may feature a strong relation or may even be used synonymously. If two or more communities – such as organisations, countries, or applications – are to be made interoperable, a significant effort may arise when each participating system is trying to understand every potential partner. In order to overcome those situations, dedicated semantic gateways are usually operated that translate, relate, and mediate the concrete representation of "meaning" between different domains of applicability and formerly incompatible services. Traditionally, a pivot representation is introduced that serves as an intermediary for all others involved, with the benefit that each systems' natural representation may only relate to the pivot.

Figure 12: Pivot Mapping



4.3 Implementation of Common Identity Patterns

Especially to support the efficient use of digital identities in distributed and networked identities patterns such as single sign-on and identity federation have evolved. This section introduces some of the most common state-of-the-art identity patterns. Patterns specific to GINI will be discussed in later sections of this document.

4.3.1 Single Sign-On (SSO)

Currently, many services and applications are offered through the World Wide Web. For guaranteeing security and protecting privacy, those applications frequently require authentication for their use. Since these services or applications are generally offered by different service providers, a user needs to authenticate at each provider separately. Taking the username/password scheme as example, a user needs to remember a single password for each service provider. Over time, this can lead to an increasing number of passwords a user has to remember. Due to that, most users tend to choose easy-to-remember passwords or to re-use one password for different service providers. This leads in a lack of security.

To overcome this issue, the concept of Single Sign-On (SSO) has been developed. Single Sign-On is defined[12] as:

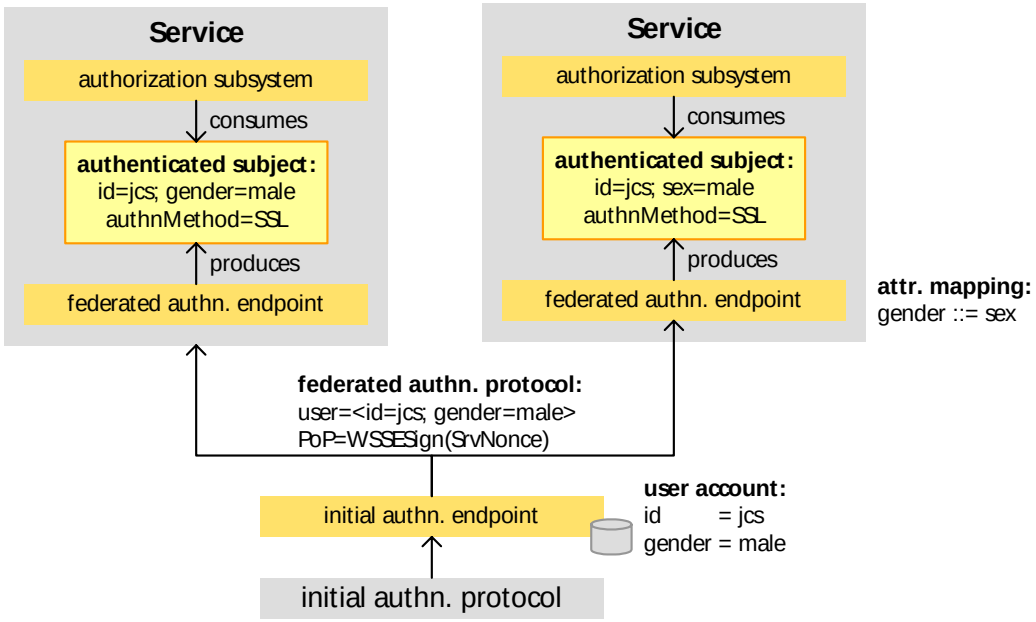
"the ability for a user to authenticate once to a single authentication authority and then access other protected resources without re-authenticating."

This means that by the help of SSO a user just needs to authenticate once in a distributed system. All other authentication processes are carried out automatically without user interaction. Going

[12] Clercq, Jan D.: Single Sign-On Architectures. In: G. Davida, O. R.(Hrsg.): Infrastructure Security: International Conference, InfraSec 2002 Bristol, UK, October 1-3, 2002. Proceedings Bd. 2437, Springer Verlag, 2002, S. 40–58.

back to the example of username/password authentication, a user needs to remember only one password with high strength which increases security.

Figure 13: Single Sign-On



4.3.2 Identity Token Linking and Chaining

A major use case for digital identities is granting rights to identified users or providing individualized services for known users. Respective scenarios need information on different aspects of the user and the usage context. Following the paradigm of a separation of duties different aspects of a digital identity and its usage context are commonly managed by different components of an identity space (e. g. attribute services).

To serve these scenarios, each identity's information providing service encapsulates its provided identity attributes as an identity token. An identity token can refer to another identity token. Multiple tokens can be linked to a single token and token can be linked in a way that they build a chain of identity tokens. The most common pattern build upon this chaining of identity token is to have an authentic token as an anchor and to link further identity tokens to this anchor.

An example of an implementation of this pattern is the European epSOS infrastructure for cross-border exchange of health data. In epSOS an identity service provider in the country of care issues an authentic identity for the healthcare professional (HCP) who is treating a foreign patient. Together with information on the HCP's profession and roles the HCP digital identity is encoded

as a SAML assertion. As the identity service provider does not have information on the treatment context, a second SAML assertion is issued by a security token service at the point of care. This assertion further classifies the treatment relationship and the treatment context (e.g. emergency treatment). By linking this assertion on the treatment context with the assertion on the HCP identity, an identity consumer (in the epSOS case this role is taken by the patient's country of affiliation which keeps the patient's medical data) gets sufficient information to decide on a medical data access request. It must be noted that the context assertion is not self-contained; it does not contain a digital identity nor is the consumer able to verify the authenticity of the assertion that the owner is the one he claims to be. This information is only contained with the original identity assertion and therefore the context assertion can only be processed if the underlying identity assertion is presented, too.

4.4 Trust and Trust Relationships

Trust is an amorphous concept. Many attempts have been undertaken to define the term 'trust', but a universally agreed definition is yet to emerge.[14] To large extent, this may be attributed to the fact that the term often receives a context- or discipline-specific connotation. In certain contexts, the term has even used to convey entirely disparate meanings.[15] Notwithstanding these discrepancies, there does appear to be a common baseline understanding of the meaning of trust; at least from a high-level perspective. In almost all disciplines, the existence of a trust relationship is typified by willingness, of one entity (the trustor), to accept vulnerability based upon positive expectations of the intentions or behaviour of another entity (the trustee).[16]

In the context of identity management, trust is typically understood in its operational sense.[17] From this perspective, an entity can be said to trust a second entity when it makes the assumption that the second entity or system will behave exactly as it expects.[18] Or when, in absence of such an assumption, it demonstrates the willingness to assume the risk associated with the transaction in spite of the absence of certainty (e.g., when relying upon validity of a credential upon proper completion of an established authentication protocol)

Trust is commonly understood to display the following features:

[14] D.M. Rousseau, S.B. Sitkin, R.S. Burt and C. Camerer, 'Not so different after all: a cross-discipline view of trust', Introduction to Special Topic Forum, *Academy of Management* Review, 1998, Vol. 23, No. 3, p. 394.

[15] See for example D. Gollman, 'Why trust is bad for security', *Electronic Notes in Theoretical Computer Science* 2006, vol. 157, 3-9.

[16] D.M. Rousseau, et al., 'Not so different after all: a cross-discipline view of trust', l.c., p. 395. See also J. Dumortier, N. Vandezande, C. Hochleitner and K. Fuglerud, 'D.7.1 Legal Requirements for Trust in the IoT', uTRUSTit Deliverable, 2011, p. 7 et seq., available at http://www.utrustit.eu.

[17] J.C. Buitelaar, M. Meints and E. Kindt (eds.) 'D16.3 Requirements for Identity Management in eGovernment', FIDIS Deliverable, 2009, p. 13, available at www.fidis.net (hereafter: 'FIDIS 16.3').

[18] Id. Definition based on Lead Study Group on Telecommunication Security, Security Compendium Part 2 - Approved ITU-T Security Definitions, available at http://www.itu.int/ITU-T/studygroups/com17/def005.doc, last consulted 10 March 2009, p. 51 and L.G. Zucker, 'Production of trust: Institutional sources of economic structure, 1840-1920', in B.M. Staw and L.L. Cummings (ed.), Research of organizational behaviour, JAI Press Inc., London, 1986, p. 53-111, and S. Slone (ed.), Identity Management. A white paper, 2004, available at http://www.opengroup.org/onlinepubs/7699959899/toc.pdf, last consulted 15, February 2009.

- subjectivity;
- evidence-driven and not self-declaratory;
- context-dependent and not consequentially transitive;
- non-symmetric;
- multi-entity, multi-agent, but not necessarily sharable
- instability, and
- which may be partially refined by reputation properties[19].

As a result, trust assessments within sophisticated systems feature constraints, such as a degree of confidence, an uncertainty factor to reflect potential risks, and a repudiation provision. In contrast to traditional systems, the degree of confidence is not only incorporating the strength of the technical means used to safeguard the trust provision but addresses all properties listed above.

Since trust may only be established between two or more entities of potentially different natures, the concrete realisation of the trust relationships between those entities may be formed in different way. The most common trust relationships are briefly introduced in the following sections.

4.4.1 Direct Trust

In a direct trust relationship, one party usually fully trusts the other party without the use of any intermediaries or other third parties. A common definition "*is when a relying party accepts as true all (or some subset of) the claims sent by the requestor*".

4.4.1 Indirect Trust

In an indirect trust relationship, the affected parties solely rely on claims asserted by a common third party with which a pre-existing trust relationship is already established. No trust path between the two communicating parties is created, since no mutual trust between the two parties is established.

4.4.2 Brokered Trust

In a brokered trust relationship, one party implicitly trusts the other partner despite having no direct trust relationship to each other by the mediation of one or more intermediaries. While the two service parties have no valid trust path between each other, the intermediaries are usually known to each other, construct the trust path, and feature a trust relationship that is at least as stable as the resulting trust relationship between the two service parties shall be.

[19] The reputation of an entity is commonly defined as: *"Reputation is the perception that an agent creates through past actions about its intentions and norms"*. (Mui, L.; Mohtashemi, M.; Halberstadt, A, A computational model of trust and reputation, System Sciences (HICSS), 2010 43rd Hawaii International Conference, 2002, http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=994181).

4.4.3 Direct Brokered Trust

In a directly-brokered trust relationship, the first party wants to engage a third party that has no valid direct trust path. The third party however directly trusts a second party that also trusts the third party and is subsequently creating a valid trust path by vouching for the third party.

4.4.4 Community Trust

In a community trust relationship, two parties create a valid trust path by their enrolment in a certain authentication community and a subsequent acceptance of its norms and practices. Apart from the communities established authentication norms and means, no additional intermediaries are introduced or used.

4.5 Conclusion

The provisions of this chapter serve as the foundation not only for this document and its contained use-cases, but also for the development of the privacy framework in D4.1 and D2.1 – "Logical Outline of the INDI Service Framework". All concepts and paradigms that have been presented in here matter crucially for the more specific deliverables in those subsequent reports.

The INDI ecosystem aims to provide a flexible, non-discriminatory, and extensible service architecture for identity services, which, in turn, requires defining certain common, stable building blocks. Each such entity will be further refined and measured against the requirements of specific functionality in the "Logical Outline of the INDI Service Framework". The trust relationships, which have also been presented in this chapter, are of special importance as they justify and sanction all interactions between all INDI entities.

In the ensuing chapter, technologies, patterns, and concepts are related to real world perspective through the sketching of concrete use-cases and application scenarios.

5 The INDI environment

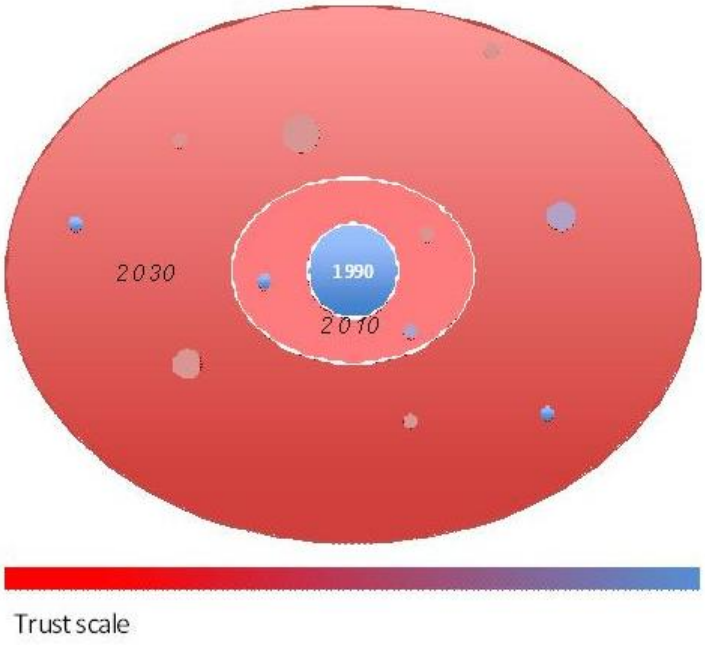
5.1 High level Gap Analysis

Today's world is marked by the phenomenal but unwieldy expansion of networks for digital transactions. While above all dominated by the Internet, there are now also "social networks" and other mediums which add to the rapidly evolving digital universe. Notwithstanding the presence of massive benefits and opportunities following from this development, conditions that would allow for orderly authentication of data or subjects, ensure effective accountability, and enable desired levels of privacy and integrity, are largely lacking. On this basis, users have limited means to influence what levels of "trust" to opt for when engaging in digital communications.

Obviously, there are partial exceptions. Particular public networks, encrypted communication, secure servers, etc. have been put in place to support stronger authentication and security within controlled spheres of transactions.

On this basis, Figure 14 presents a stylized illustration of our present world, and how it compares to what was there in the past, as well as with what may follow. Here, the scale applied shows, in the one extreme, "unsecure" communication and transactions marked in red. At the other end, those that are "fully" authenticated and trusted, are marked in blue. In practice, a particular technology and market situation would not show up in any of these extreme corners. In between there a gradual colour shifting from red to blue, which marks to what extent conditions lean in one way or the other.

Figure 14: Stylized states of trust in digital transactions; past, current and extrapolated (1990-2030)



The present situation, marked by rapidly expanding data bases and communication under conditions largely lacking the tools to ensure security and trust, contrasts sharply with what was there in the early 1990s. At that time, denoted by the smaller round sphere at the centre of the figure, digital transactions were obviously much more limited in scope. Many of the (relatively few) people and organisation's exchanging electronic information were in close contact with one another

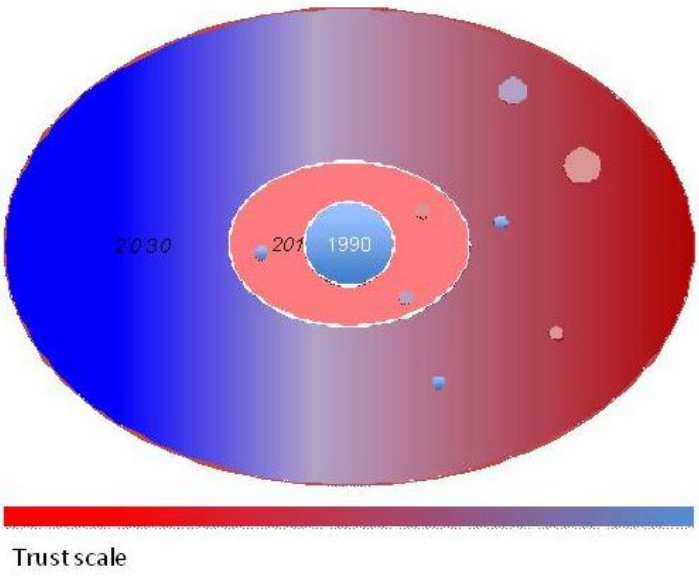
through other means of communication anyway. That was also how the Internet was essentially started, as a tool for exchanges within a network of researchers who basically knew one another already. Hence, applying the colour scale used here, it can be said we belonged in a state as illustrated by the smaller uniform bluish sphere at the centre.

Again, as for today, most digital communication takes place under largely insecure conditions, reflected in the relatively more reddish space overall compared to 1990, although there are also those "islands" of more secure exchanges.

As we look ahead, our present largely insecure digital universe may evolve in different directions, as illustrated by the diverse outer rings in Figures 14 and 15 respectively. If things continue as now, extrapolating the current trends implies ending up with a situation in, say, 2030 as illustrated in Figure 14. Most of the exchanges would continue to be unauthenticated and insecure. For that much larger digital universe, it would be fair to assume there would be dire consequences overall. As a response, a range of fragmented counter-measures would evolve, here showing us as secluded networks and spheres marked by strong authentication, depending on what "state-of-the-art" defence, may enable shielding particular kinds of clients against the "ordinary jungle". Hence, the bluish circles have become much more frequent, and some of them larger and more developed, than what we have today. Some would be run by public interests, others probably by private ones, which would be in limited areas where the suppliers of such identity management frameworks could appropriate adequate financial returns from their service.

There is a possible alternative scenario for 2030, as illustrated by Figure 15. This is one in which differentiated levels of trust evolve, in an orderly manner, allowing for informed, customer-driven choices and differentiation. This is the state of play aimed for in the GINI project. In this case, users would be able to express their needs for identity management and associated services, and there would be multiple operators available to service their INDI environment, on the basis of their particular needs and preferences, when it comes to levels of privacy, protection, insurance, or "levels of trust". There is, indeed, a rationale for varying kinds of services. Scientists operating in different disciplines, and working with different kinds of data, industrialists, or the general public, would not require the same level of trust in managing their identified, or in accessing and exploiting data.

Figure 15: Stylized states of trust in digital transactions; past, current and those of a differentiated future (1990-2030)



For informed choices to be made, there will have to be an evolution of authentication service providers. These would specialise in accordance with what they work out to be their "core business", and so as to be able to gain the capabilities of developing, offering and tailoring the best possible solutions with consideration paid to the preferences displayed by different kinds of users and customers.

As implied by Figure 15, whereas a great deal of communication would probably still be based on weaker forms of identity management, and lower levels of trust, in this kind of world there would be a wide spectrum of opportunities for requesting and deploying high value-added authentication services as well. There can be little doubt that the latter state represents the more desirable future state of affairs.

5.2 Primary Motivation: Privacy Enhancement

For some time now, there has been a growing tension between the collection and storage of personal identity data in digital transactions and the respect for users' privacy. Although federated identity shows potential to mend several of the issues related to digital transactions, security and privacy concerns still continues to hinder its progress.

The call for counteraction implies that the Internet will not continue as we know it today. Whereas governments will act to secure areas of highest priority to them, and informed and established private actors can be expected to protect their data and transactions, there will be the risk of sharp segregation between those who are informed and protected, and those who are floating in the jungle.

Faced with such risks, governments have a key responsibility to take the actions that are necessary to ensure the development of an orderly future universe for digital communication. Given the cross-border nature of the digital world, governments must also collaborate closely internationally, to develop effective and affordable responses. Again, however, this does not imply that public sector actors will be the best placed to develop and implement the solutions of the future.

In this context, GINI will examine the technological, legal, regulatory and privacy-related dimensions of the gap between the current state of the art and the vision for an INDI ecosystem beyond 2020. The overall framework will account for a market space capable of dynamically developing services for the handling and storage of identity data, based on the proportionality and minimization principles.

Putting People in Control

GINI strives to establish a Personalized Identity Management ecosystem by bringing in different stakeholders from public institutions, businesses and civil society to set regulatory and operational framework for operators supplying INDIs within the EU. The framework will be an umbrella to establish and control INDI operators to perform their services in a regulated environment. Those operators will operate in a competitive business model where new businesses will evolve over time, including existing identity operators such as banks and telecom.

GINI objective: In practice individuals only have limited control and knowledge on how and where identity data is collected, stored and processed. It is the objective of GINI to outline a digital identity ecosystem that puts individuals in maximum control of their digital identities.

5.3 The INDI ecosystem

5.3.1 The INDI as a User-centric digital identity

An Individual Digital Identity (INDI) is an identity claimed in the digital world by an individual who creates, manages and uses it. Individuals will have the ability to establish and manage an INDI and decide where and when to use it while interacting with other persons or entities. As a result, individuals will be able to present their chosen, verified partial digital identity to other individuals or relying parties with which they wish to build trust relationships in order to perform transactions for personal, business or official purposes.

The INDI is a digital identity that is:

- Self-created by the individual
- Self-managed throughout its lifecycle (creation, change, management, revocation etc.)
  - Either with IT system support in the domain of the individual
  - Or through the assistance and support of an Operator under a service model
- Verifiable
  - Against authoritative registers or data sources that the user selects
  - Only when, and to the degree that, the user chooses
- Presented to entities with which the individual enters into agreements and service transactions
- Presented to other individuals with which the individual conducts online transactions and/or communicates

5.3.2 A Network of INDI Operators

The INDI ecosystem is based on a network of "Operators". The rationale to choose an operator network model as the basis of a user-centric ecosystem is as follows:

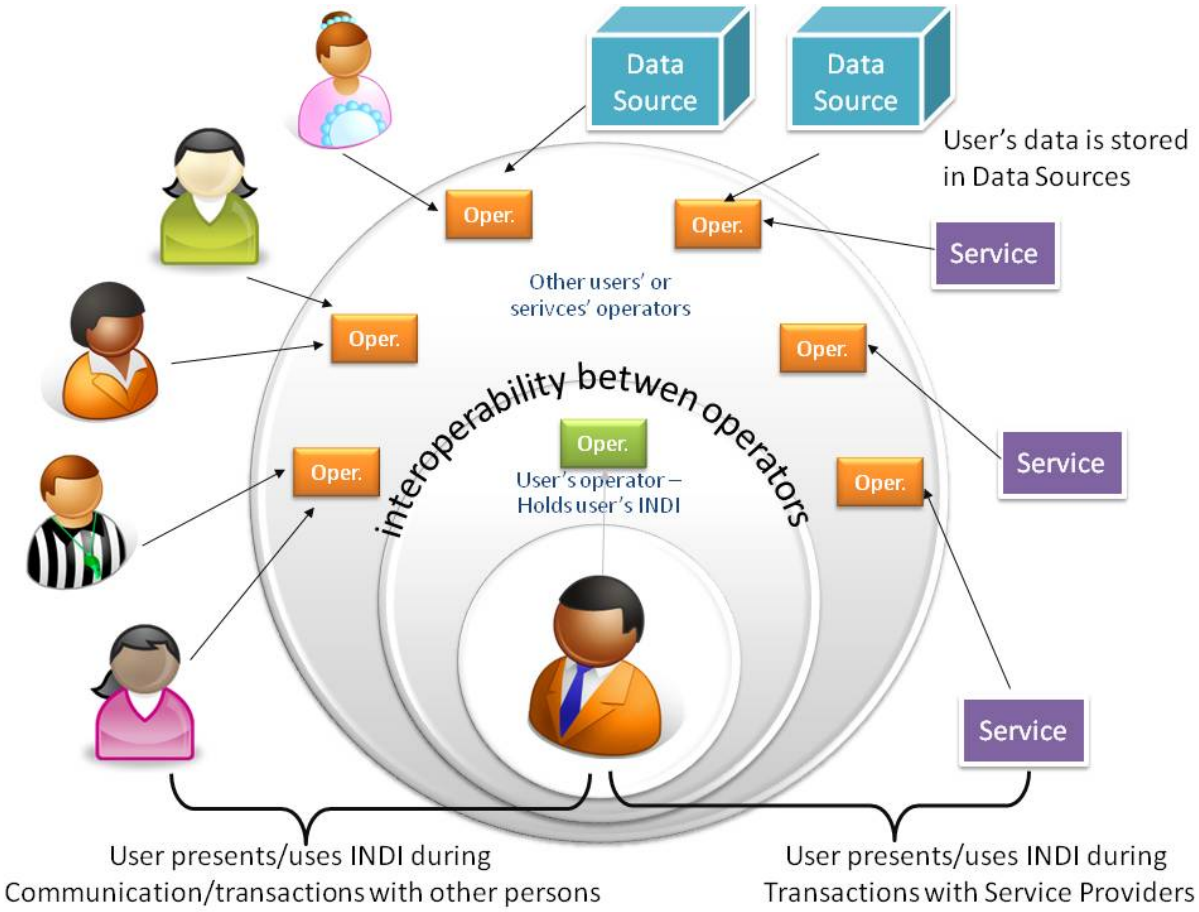
- Independent Trust Anchors are needed to enable trust within the INDI environment and provide added value beyond users' self-asserted claims
- From a risk management and privacy point of view, it is important to avoid centralised single points of failure which also present threats for privacy-compromising data aggregation and/or profiling; INDI management and data should be de-centralised and decoupled from each other
- The INDI Operator concept and the business models it enables make it easier to create a truly global and competitive market for INDI services
- Users cannot manage trust decisions, 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 – users should have sufficient technical assurances and legal warranties so that their "trust decisions" can then be safely based on the usage

and services of the INDI Operator network, as represented by the INDI Operator without having to question every entity of the underlying infrastructure.

- The Operator Network model can be standardised and regulated easier than a model, which is based on very heterogeneous and un-even entities, and this can greatly enhance the users' ability to build trust relationships with Operators.

A high-level view of the relationships formed within the INDI ecosystem follows:

Figure 56: Relationships within the INDI Ecosystem

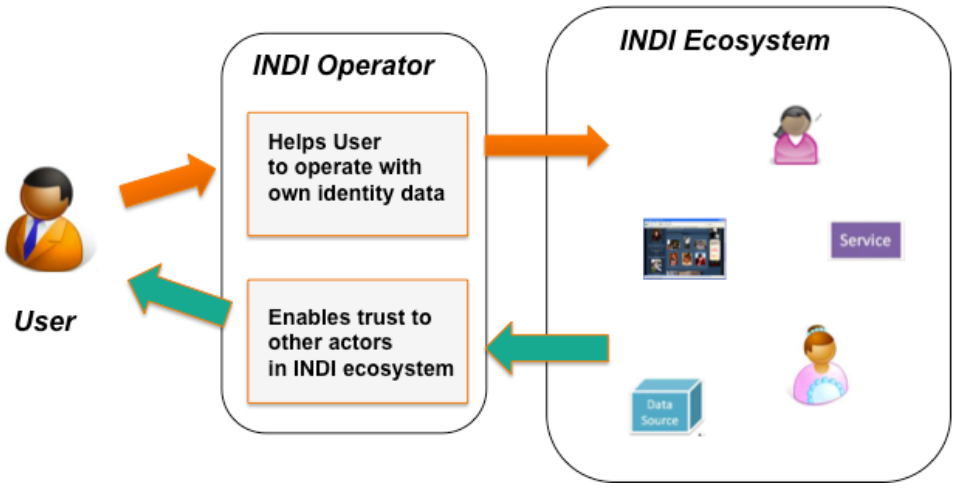


The INDI allows the individual to act in various roles, for instance citizen, employee, customer, either connecting them to one INDI or having multiple INDIs for different interactions. The user chooses which roles to act in and what information to reveal in the different roles. As such, a INDI serves to represent the user in many different contexts. However, the user is able to manage its partial identities similarly as in the physical world, by providing the relevant information to each situation, including cases where anonymity, pseudonymity and limited attribute provision would be desired and acceptable

When individuals act in a delegated role on behalf of another person or a legal entity, certain rights and responsibilities may be linked temporarily to an acting individual's INDI as a mandate or role delegation, with the possibility to be de-linked at the end of the desired period when this functionality should be allowed. On the other hand, permanent (but revocable) delegation and mandates could be supported, which official authorities could recognize.

5.3.3 The INDI Operator and the User

Figure 67: The INDI Operator and the User



The most important relationship within the INDI ecosystem is the one between the INDI Operator and the User (the INDI holder). Although the INDI ecosystem does not exist yet and, when created, it will stretch beyond of what is happening today, it is easy to imagine that the INDI relationship has similarities to the existing international operator models from other sectors outside identity management:

- In the banking network, the User establishes a relationship with a bank, which connects the user to the network of banks. This relationship makes it possible to transfer money to the users or organisations who have relationships with other banks.
- In the telephone network, the User establishes a relationship with an operator, which connects the user to the network of operators. This relationship makes it possible to make calls to the Users who have relationships with other operators.
- On the internet, the User establishes a relationship with a network access providers (in fact, with several ones) and in every such agreement after access is granted an UP is assigned which enables the User to access everything on the Internet, of course always subject to access policies of content and service providers.

The key question of the INDI ecosystem is what does a relationship with an INDI Operator mean for the User? Without getting into the functional details, the relationship can be described as a trust relationship related to the identity data. An INDI Operator enables the User to use the User's own identity data with the users or organisations who have relationships with other INDI operators.

The trust relationship between the INDI Operator and the User consists of two different parts:

- The INDI Operator can be a trust anchor, which helps to verify the User's identity data in the INDI ecosystem – the whole ecosystem has a trust relationship with the user through the INDI Operator

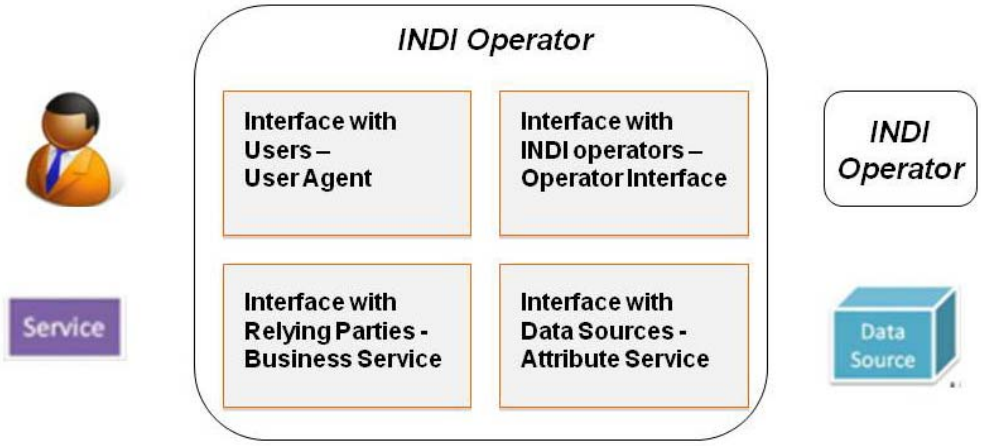
- The INDI Operator is a trust anchor, which helps the user to trust the other actors in the INDI ecosystem – other users, organisations providing services and data sources

The following additional notes can be made about the relationship between the INDI operator and the User:

- The INDI ecosystem is global, which means that the INDI Operator and the User need not to be from the same country or identity domain.
- The relationship has a contractual and legal dimension and not just a technical side.
- The User should be able to have relationships with several INDI operators at the same time and in parallel, and also be able to switch from one to another, much as can happen with mobile telcos.

5.3.4 External Interfaces of the INDI Operator towards INDI actors

Figure 78: The INDI Operator and Other Actors



The INDI operator has four possible external interfaces, which define the architectural boundaries, where the INDI Operator operates.

INDI Users operate with INDI Operators with help of a User Agent, which can be an application or a service. The User Agent interface has the following requirements:

- Users must be able to get access to their identity data.
- Users must be able to communicate with services or other users.
- Users must be able to manage consent to give data to others in the ecosystem.
- Users must have access to their operator account information.

Relying Parties and Data Sources can access or be accessed from the INDI ecosystem using Operator-side interfaces which are called Business Services and Attribute Services respectively. Data Sources Business Services and Attribute Services have the following requirements:

- Interfaces enable INDI ecosystem requests, which originate from the other actors.

- If access to a service or data source is chargeable, the interface must include a method to receive payments.

Finally, INDI operators are connected to other INDI operators, who together form the INDI ecosystem and INDI operator network. The INDI Operator interface has the following requirements:

- The interface must enable requests to services and data sources, which are facilitated by the other INDI operators.
- The inter-Operator interface must serve requests, which are originating from, and targeted to, other User Agents, Business Services and Attribute Services provided by the communicating Operators.

It should therefore be considered that User Agents, Business Services and Attribute Services are all Operator interfaces on the side of the three main actors of the INDI ecosystem (the User, the Relying Party and the Data Source). In principle, business models for Operators will emerge, which offer a combination of all or some of the above interfaces; it will be a matter of further investigation in GINI to determine under which conditions these interfaces can co-exist in the same Operator, provided privacy requirements are met.

These considerations will be in the scope of D3.2 and the updated version of D4.1 (M16); a more detailed architectural description of these Operator interfaces can be found in D2.1.

5.3.5 Principles of Data Disclosure within the INDI ecosystem

In principle, there are two options for verification of an INDI:

a. The user submits data to the Operator and these are verified against data sources of the individual's choice. This implies the following:
   - That the user relies upon the Operator, although there is a significant disadvantage of possible data aggregation at the Operator.
   - This risk can be spread by using more than one operator for different identity domains, at the expense of less usability in the absence of a one-stop service point.
   - Interoperability between operators and INDIs is a pre-requisite of the envisaged INDI infrastructure, allowing for INDI portability and INDI dispersion (i.e. spread between operators).
   - When an INDI is presented to/used by an individual, or a Relying Party with which the INDI user wishes to enter into an online transaction, any request of data can be handled by the Operator, but only after explicit consent of the INDI user.
   - Assertions to the authenticity of data can be issued by the data source directly to the recipient, or through the Operator. In the latter case, the Operators also aggregate more information on transactions (which is a disadvantage), but the Data Sources do not need to implement a more complex service interface.

b. The user does not submit data to the Operator but points to the data source where the data is located, and registers verified (and verifiable) links to those data. This implies the following:

- The INDI user makes use of some token issued by the data source to register verified links to data with the Operator.
- The Data Source issues assertions that the data in question are available and valid, but does not disclose the data unless under specific circumstances to which the user has consented either in advance or on the fly. Disclosure may occur without explicit user consent in cases mandated by law (e.g. fraud, etc., cases when privacy can be overruled).
- When an INDI is presented to/used by an individual or a Relying Party with which the INDI user wishes to enter into an online transaction, any request of data will be handled by the Data Source directly to the recipient party. If an Operator is used by an individual or a Relying Party that is the recipient of an INDI, then a similar policy of non-aggregation of data will have to be enforceable. Operators could be regulated to support both models.
- This option has the advantage of avoiding centralized data aggregation and should be preferred. But it also comes with preconditions of technological interfaces and trust models that go beyond the current state-of-art.
- Operators under this scenario will have more of a routing functionality and will act more as intermediary trust anchors offering ease of use. Different business models not based on data aggregations should be explored.
- Data sources would have to implement more complex service interfaces than with option a.

In reality, the User might be able to choose between the two options above and authorize partial disclosure of data. In this context, current (and future) technologies for minimal data disclosure will be very useful and their use could be made mandatory. Such technologies can be used on the side of the data source in order to give to the user the possibility to manage data disclosure to the Operator, but they can also be used from the side of the Operator so that the User can manage data disclosure to INDI recipients.

Moreover, an initial verification with a data source might take place at the moment of creating an INDI, i.e. before the INDI is presented to any party or otherwise used. An Operator then would keep a record of assertions resulting from such verification, together with reference to the trusted data source, which provided them.

Except for the initial identification process to establish an INDI by linking it to verifiable and authoritative data sources, it should be possible for most (if not all) other transactions to be handled through pseudonymity, ensuring the privacy of the User. Thus, when a User is interacting using an INDI, it is not possible for anybody to know the identity of the User, unless the user opts to divulge its identity or in cases mandated by law. This also puts certain responsibilities on the user and transfers to the Relying Party the choice on whether to accept the User's wish or enter into a negotiation. This is expected to be a welcome departure from the present situation where a consumer is forced to either accept the Relying Parties data disclosure wishes in order to make use of their services, without given any option to disclose fewer or different items of data instead.

5.4 The INDI as a User-centric Address

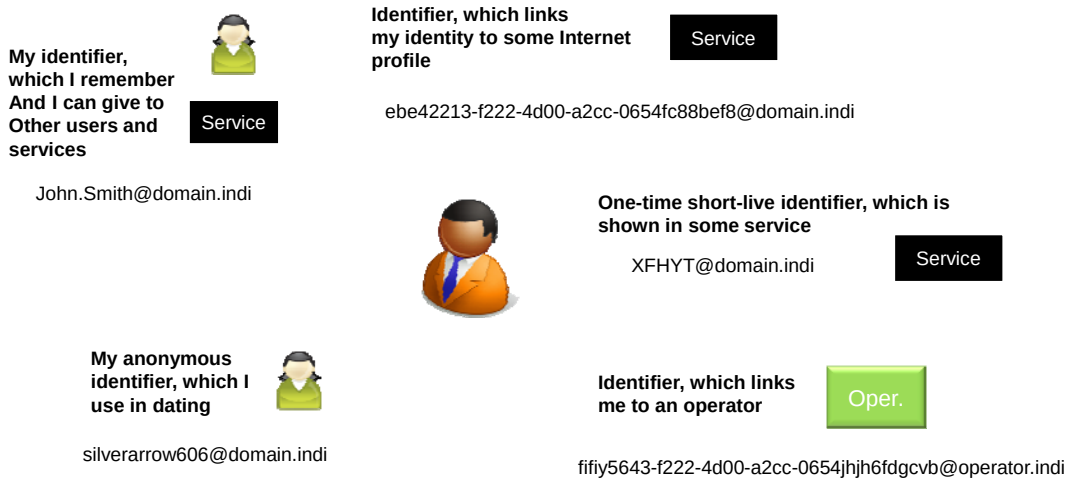
An INDI address can be presented to a Relying Party and can be connected/referenced to an INDI (either directly or indirectly) within one or more domains inside the multi-domain INDI environment. An INDI domain in this context can be considered any an application or service area where an individual uses a specific INDI as a domain address

The INDI address consists of two parts:

- Domain part, which describes the particular INDI domain where the user wants to use an INDI as an address, and which the user somehow controls.
- Address identifier part, which describes an INDI representation within the INDI domain.

No specific syntax for an INDI address needs to be defined at this stage, but for conceptualization purposes an e-mail address type syntax is used here as an example (indi_id@domain.indi).

Figure 89: The User-centric INDI



INDI addresses are needed for several use scenarios. Because of different needs, several requirements for the INDI addresses can be set:

- One user can have several INDI addresses.
- It should be easy to generate and change INDI addresses.
- It should be possible to generate INDI addresses, which are specific to some particular service.
- It should be possible to generate INDI addresses, which are short-live by nature.
- INDI addresses should not be operator specific.

5.4.1 INDI address, which the user can remember

Users need some e-mail type address, which they can remember when the address is asked by another user or a service. The e-mail type address can be used to initiate INDI communication with the user. One User can have several addresses and it should be possible to change the address easily.

5.4.2 INDI address linked to an Internet profile

If the User wants to link his INDI to an Internet profile (such as user profile in an Internet community), a service specific INDI address might be needed. If the address works only in the context of some particular Internet community, the communities and users do not have to worry so much about the security of the address, because it is unusable outside the community. Additionally, if the user changes the INDI address description (i.e. a handle), the addresses in different communities need not to be changed, whilst on the other hand a known handle that serves as INDI address description might be transferrable even when the User changes Operators; this can allow INDI portability.

5.4.3 Anonymous INDI address

Anonymous INDI address is similar to the one that the User remembers, but it is intended to anonymous use. Because of this, the address should not reveal anything about the User.

5.4.4 One-time short-lived INDI address

There might be a need to present an INDI address to the User in public Internet (for example in advertisements, user profiles or e-mail). For improved security, it should be easy to generate one-time, short-lived INDI addresses (e.g. valid for one day and which can be used from one network address only).

5.4.5 Operator-specific INDI address

In an INDI environment, the User has a connection to the Operator and the Operator may give the INDI user some INDI address. However, it is very important that the INDI addresses are not tied to Operator domains because it should be easy to change the Operator.

5.5 Using an INDI

5.5.1 Presentation of own Verified Data to Individuals or Relying Parties on the Internet

The INDI environment allows Users to present their INDI towards other physical persons or legal entities with which they wish to build trust relationships in order to perform transactions for personal, business or official purposes. The INDI environment should allow Users to transition seamlessly, both ways, between the physical and digital world(s).

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).

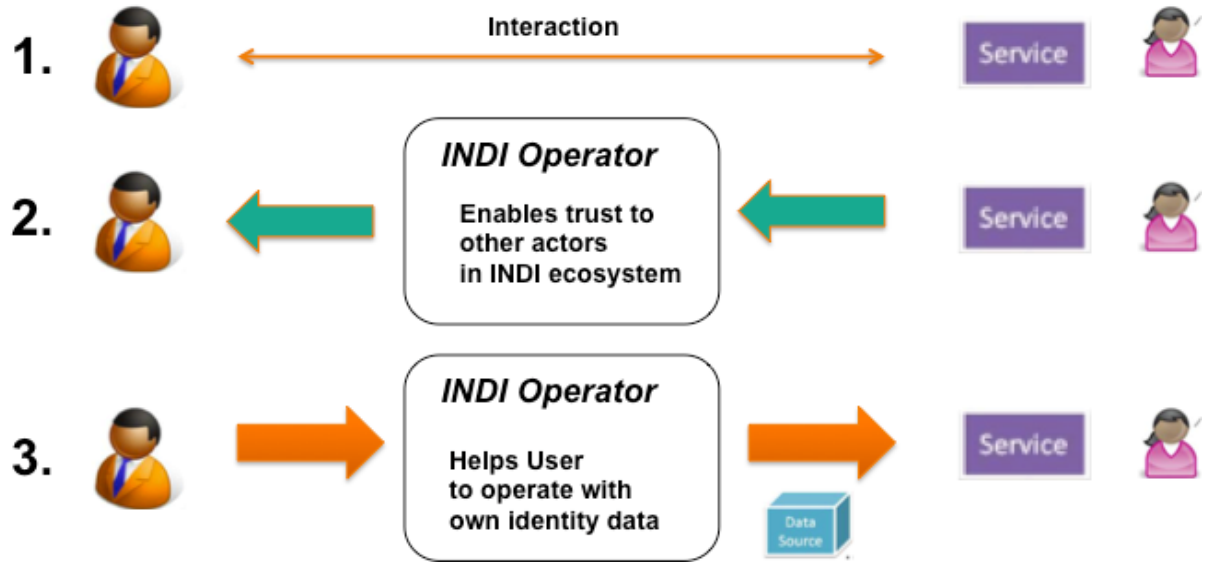
The necessity of showing data might be related to various requirements, such as authority requests (identity card, driver's license, student card) or to improve own credibility (age, place of residence, picture). The need to show a User's own data on the Internet can be met during on on-line transaction (such as registration with or use of a public service) or in on-line messaging (e-mail, chatting). Also, data might have to be shown in a face-to-face situation or in relation to an off-line profile of the person (such as a profile in a classified ads site or social networks).

First, the presentation of the data must somehow be initiated. If the User communicates with the receiver of the data in an untrusted environment, the User can present some INDI-related identifier, which can be used to create and INDI request.

Second, the User must build trust to the receiver of the data (this is the main challenge of the web pages or telephone numbers – the User cannot really be sure, who he or she is communicating with). In some cases, the User would not necessarily be interested to know about the viewers of the data.

Thirdly, the User must give permission to the INDI ecosystem to show the data to the receiver. With help of the permission, the ecosystem can utilise a trusted data source to present verified data to the receiver.

Figure 20: Interactions when using an INDI



One important aspect that needs to be defined is the storage of the data. Although this should not be an absolute constraint because Users might accept it, a basic initial assumption would be not to store data at the INDI Operator, because: reasons:

- From the security point of view, it is not good if a centralised register is created, where rich combinations of different data of the User is stored – the register would be a potential single point of failure and a target for identity thefts.
- Most data registers are by nature on-line registers, which means that the data should be refreshed anyway from the data source – some cache mechanism could be considered for use cases, where on-line connection is not available at point of the use of the data.

There are also many other issues, which must be solved within the INDI ecosystem:

- What is a driver's license (or some other data) from the User's point of view? – document, application, set of attributes?
- Does the user have to pay for something?
- What does the trusted data look like to me and to the viewer?

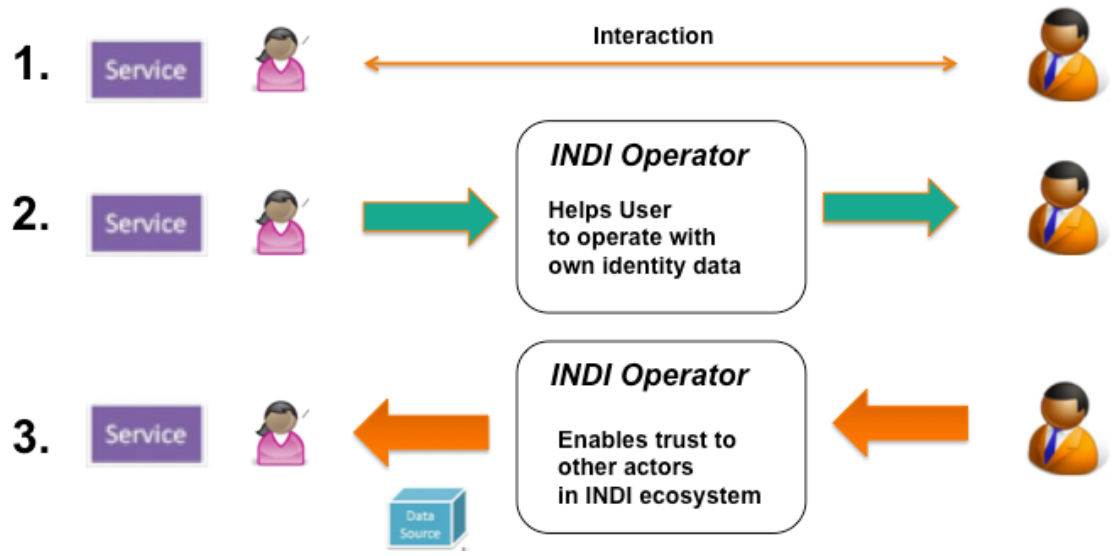
  - What should be standardised in visualization?
- How is the trust chain built from the User's perspective?
  - How can get my data in a trustworthy form?
- How can I know, who I can present my trusted data?

5.5.2 Verification of the identity data from INDI users

There are many situations on the Internet, where a User or a service would like to verify some data of some other Internet user. Examples of such cases are government processes, where an individual or Relying Party acts to check the data, selling and buying things, recruiting services, dating services, financial services, etc.

Verification of the data is the mirror of the presentation of the data and it can also be divided into three parts:

Figure 21: Verification of the INDI identity data



First, in order to check anything, some data or other link must be received from the User, which helps to connect the User to the INDI environment. The data can be document, INDI identifier, biometric data etc.

Second, the User or Relying Party must be known to the INDI ecosystem so that the owner of the data can trust the receiver.

Third, the result of the request is received from the INDI Operator (who received it from some trusted data source) and the data can be processed further.

From the requestor point of view, the credibility of the data depends on two independent sources:

- An INDI Operator (or network) ensures the integrity of the data and proper authentication of the different parties and data sources.
- The quality of the data from the data source is adequately ensured.

Another aspect, which the requestor needs to consider, is the age of the data. Basically, the INDI network could offer everything on-line, but that is probably not necessary for all applications. If non-on-line data is used, the receiver must decide, how old data can be trusted. This justification may vary greatly between the different applications. The problem can also be solved by giving an option for an on-line refresh (for a fee for example).

There are some other issues, which the INDI network must solve. For instance, how is the trust chain visualised? Hold old may the presented data be? Who takes responsibility for verification? These will be returned to in the user-cases.

5.5.3 Linking INDIs with authoritative Ids

National e-ids can be linked to the INDI and function as a method of authentication in the relevant digital interactions. Presumably, e-ids will no longer be considered to be the identity of the individual user, like some claim, but an identifier linked to the INDI, the claimed identity in the digital sphere.

5.5.4 INDIs in the Cloud

Main identified pitfalls associated with Cloud Computing:

- Lack of interoperability: due to lack of common standards, procedures, and tools, the User cannot migrate its data and services from one provider to another, which in turn causes dependency on a particular cloud provider (so called lock-in effect).
- Data protection: The customer (in their role as data controller) may lack the ability to effectively check the data handling practices of the cloud provider and thus to be sure that the data is handled in a lawful manner. The problem is exacerbated in cases with transfers of data, for example between federated clouds (e.g. Google and Amazon sharing individuals' information).

Further practices on how online services linked to clouds compromise the control of identity:

- Individuals are in general not aware of the risks associated with sharing extensive information about themselves on social networking or other sites (e.g. Google).
- The advent of companies operating by taking people's personal information from publicly available sources – such as the databases, electoral registers, company registers, phonebooks and websites – and aggregating these sources to form extensive personal data files.

The INDI environment can be used to verify the authenticity of an Operator providing clouds and cloud services. Operators must implement protection measures that ensure that no one but the individual is able to exercise control over its personal data, unless requested by the individual. The INDI environment will also require Operators offering clouds and cloud services to be transparent about the usage and handling of personal data.

5.6 The INDI Lifecycle

5.6.1 Creating an INDI

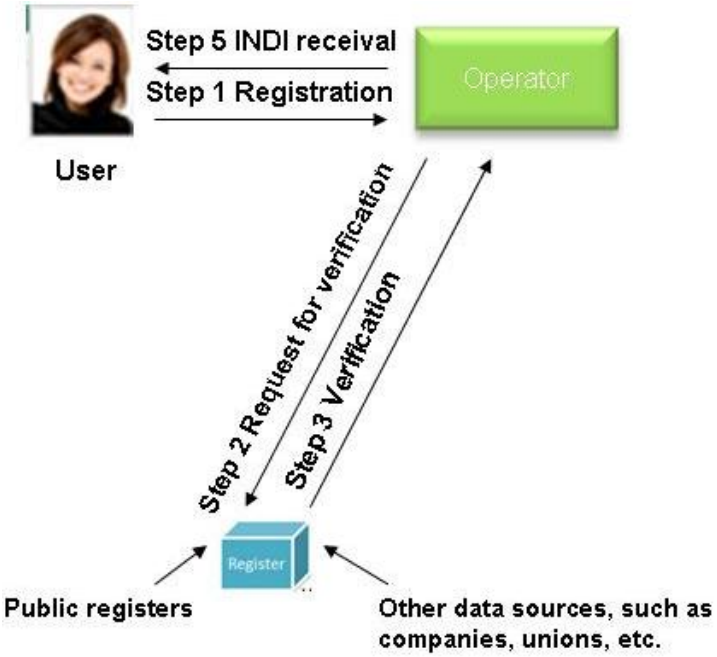
5.6.1.1 Registering a User with an Operator

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.

In order to create and use an INDI, an individual must establish and maintain a relationship with at least one INDI Operator. 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 fulfil its function, the identity must be created and enrolled with the operator service. Depicted in Figures 22 and 23 is the registration process for an individual obtain an INDI from an operator.

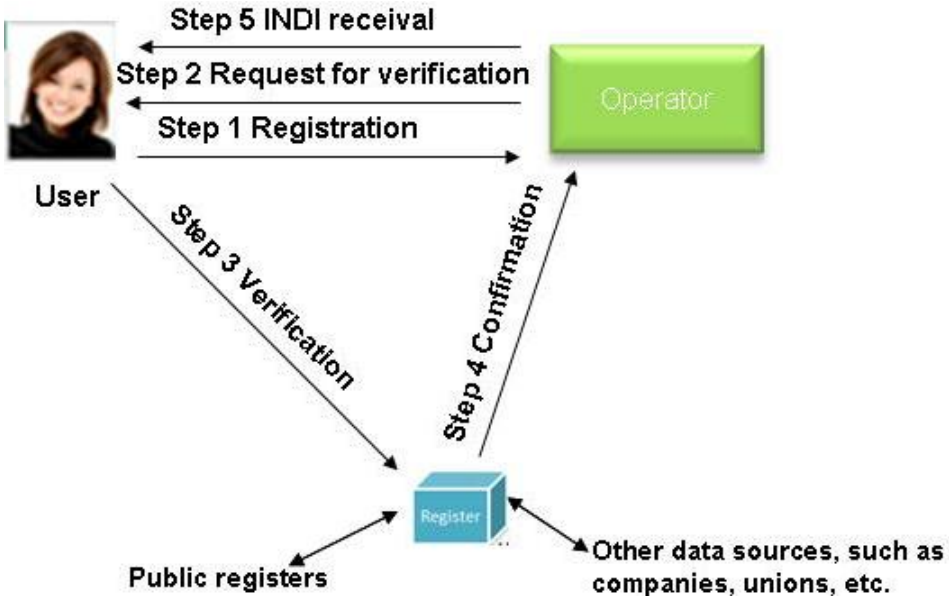
In Figure 22, the user submits data to the Operator which are verified against the data sources that have been chosen by the individual. This option requires the Data Source to have exposed appropriate interfaces and a pre-existing trust relationship with the particular Operator. Apart from giving an explicit consent, the user need not interact directly with the Data Source.

Figure 22: Data-supplying Registration to an INDI Operator



In Figure 23, the User does not submit data to the Operator but points to the data source where the data is located, and registers verified (and verifiable) links to that data. Subsequently, the User can request that the register provides the links to the Operator or send them directly. This model requires the data source to have exposed appropriate interfaces and the User to authorize (or have authorized) the Data Source that the particular Operator is trusted, but also to directly interact with the Data Source at registration time.

Figure 23: Data-linking registration to an INDI Operator



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 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.

The main services (tasks) of an INDI Operator include:

- creating and distributing INDIs and addresses for the user in such a way that they can be used within the INDI environment and presented towards relying parties;
- ensuring that no-one, including Operators, can determine the identity of the of an INDI User, unless the User divulges it or mandated by law;
- obtaining consent from Users for the disclosure of their personal information maintained by a register/data source and warrants having obtained such consent;
- enabling trust across otherwise untrusted domains: an operator may ensure and represent authentication of the parties involved (Users and Relying Parties) in a particular transaction;
- ensuring the authenticity (i.e. source and integrity) (but not reliability) of data presented towards Relying Parties;
- managing a reference directory (about which registers maintain information about Users, e.g. population centre or city resident register);
- ensuring technical interoperability among the parties involved in a transaction in the INDI environment.

The Operator in principle does NOT:

- make any warranties with regards the reliability of the content of information maintained by a registry;
- store or aggregate any information exchanged between users and relying parties.

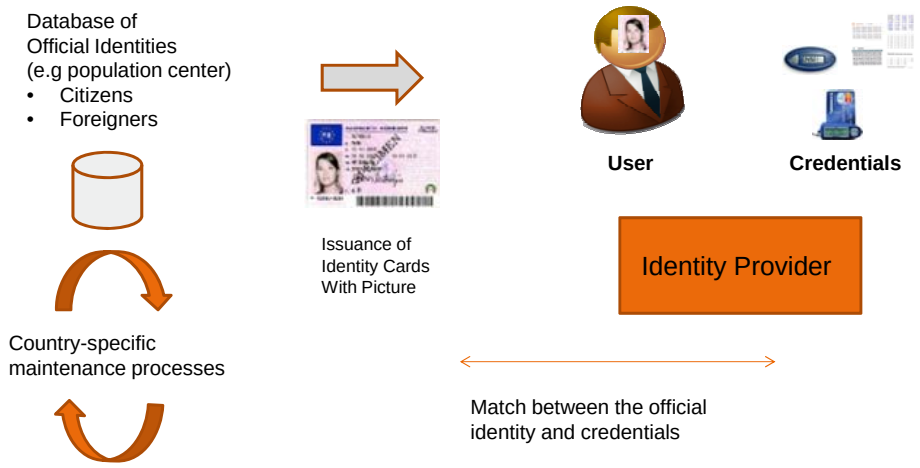
A given Operator can have a relationship with any actor in the INDI environment (users, Relying Parties, Data Sources), yet the specific terms that govern their relationship will be different In principle, interactions between Users, Relying Parties and Data Sources will be done through their respective Operators, thereby reducing or even eliminating the possibility for any given Operator to aggregate information on parties with which the user engages in transactions. In cases however where the actors use the same Operators, care should be taken that the transacting parties may not be directly linkable in a way that might be vulnerable to profiling.

The Operator model will lead to the establishment of and INDI-enabled and enabling infrastructure that will be supported by internet-world Operators. Setup principles for such infrastructure should enable interoperability between Operators, allowing for INDI portability and INDI dispersion (i.e. between operators).

In the existing digital world, there are many different kinds of practises, how digital identities are created. In the following sections we will describe three different models are presented here and analysed briefly. The list is not complete and by default, all models (and new ones) should be supported in the INDI network (unless there is some particular reason, why some method or practise not be allowed)

5.6.1.2 Enrolment, which relies on official identity and user credentials

Figure 24: Enrolment based on an existing id



In many government driven identity management projects, the basic assumption is that the authorities create and maintain the official identity attributes linked to the physical identity of a citizen. Typically, the management of official identity attributes is domestic and citizen-centric and there are one or more official registers (e.g. population centre or city resident register), where the official identity attributes of one country are defined and maintained. Normally, non-citizens can also be registered in the registers so that they can get the government id (e.g. social security number), which is necessary, for example, for living and working in some particular country.

In the context of government defined official identity attributes, the digital identity is often implemented in such a way that an Identity Provider (IdP) links authentication or signing credentials to the physical (official) identity. In practise, this means that the IdP checks a valid identity doc-

ument, when the credentials are given to the person, creating a strong link between the digital identity and physical identity.

There are several benefits from mirroring the identity of the physical world in the digital identity:

- Users understand easily the link between the physical and the digital identity.
- Identities are well-defined in the legislation.
- Unique government IDs make it easy to match identities in different services.
- Enrolment process, which is based on the physical check of the official identity document, is secure.

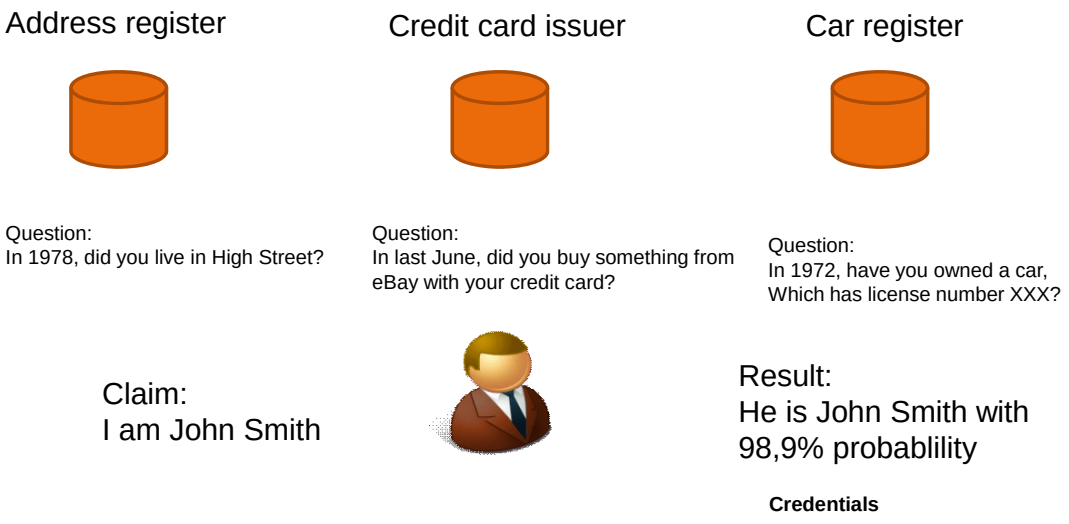
Use of the physical identity as a basis for the digital identity also meets with several challenges:

- Solutions are typically domestic and often vary between the countries, although common attributes such as passports are now more or less standardised.
- Solutions are often based on one technology, which the government defines – this has led to solutions which are difficult to use and lead to low user adoption.
- Same individual's entry in the registers of two countries leads to two different identities.
- Solutions are often more secure than most of the services require and do not provide sufficient interfaces (APIs).
- Often, solutions do not include option for anonymity.

5.6.1.3 Enrolment, which relies on user knowledge

Verification of a person's identity can be done in such a way that the user makes a claim of the identity and then some evidence is searched from data sources that the claimed identity exists.

Figure 25: Enrolment based on a verified claimed id



If evidence is found, the User must answer some yes/no questions based on the details of the data in the data sources. The more details the User knows, the higher the probability that the claimed identity is authentic.

In practise this method works in such a way that the User first provides basic identity data, such as name, birth date and address. After that, the information found in the registers about the user's address history, education history, work history, financial transaction history, legal history etc., is utilised for verification questions. The process can be modified based on the context and several sets of questions can be used. As a result of the process, a probability is calculated and a business decision is made based on the results obtained.

A user-knowledge based method is often used in countries, which do not maintain centralised population registries, such as in the US and in the UK. A process can be used to issue credentials or it can replace credentials in such cases which are not frequent (e.g. issuance of a credit card). The method can be used together with many technologies and it can be modified based on the registers that are available.

A user-knowledge based method has several benefits:

- The method is practical, easy to understand and user friendly and does not include technology.
- The method is flexible and can be customised based on the context.

Obviously, the method has several also challenges:

- Quality of the method is based on the availability and quality of data registers.
- It offers a low level of assurance (e.g., if somebody collects enough data from users, they can pass the test questions).
- Users are not confident about giving their data or answering questions on the Internet.

5.6.2 INDI Issuing/Building

An INDI can be obtained from Operators within the INDI ecosystem. To start using the INDI, the User needs to enter minimum set of details. It is vital that the User details do not reveal the true identity of the user, unless the User voluntarily reveals such information.

The User can set an optional number of privacy policies for using the INDI. Examples of theses can be:

- Accessing an INDI, e.g., security policy, back-up policy, etc.
- Using an INDI in an interaction.
- Signing by using an INDI.

The User of an INDI can set up his chosen method of authentication as a private policy. The preference of authentication features depends on how much or how little security and privacy the User may choose for regulating the access to the INDI. As the User may have INDIs for different contexts or roles, the level of security for each INDI is to equal the personal risk appraisal made by the INDI User for different transactions against the security level set by the Relying Party. Typically, the individual can select available premade sets of policies; define the credential, which only the User may know. Although the packages are premade, nobody but the User of the INDI knows which package or credentials that have been chosen.

However, if the individual User is acting in a delegated role for a legal entity, the legal entity will create the policies that it wants those in delegated roles to use. As part of the assignment of a delegated role, the User of an INDI inherits these policies.

5.6.3 Revoking an INDI

To revoke an INDI requires that the User terminates its associations with its Operators or simply stops using the INDI in digital interactions. Still, it is possible to re-establish an association with an Operator at a later time.

5.7 INDI Administration and Protection Considerations

An Operator aims to keep the individual in maximum control of her identity data. This means that only the individual personally may:

- consent to the collection and/or processing of her identity information;
- alter/edit her identity related information;
- link and unlink identity data sources to an INDI environment;
- decide on which identity information are disclosed to whom and for which purpose by defining policies or ad hoc decisions.

The Operator must therefore implement protection measures that ensure that no one but the individual personally – or a dedicated proxy – is able to exercise control on that individual's identity data. Even more, the Operator must enforce the individual's will, as expressed by his or her specific configuration of the INDI environment.

Deletion should also be an option, although the service provider must also have the responsibility of deleting personal data after the legally defined retention period is passed, or in other circumstances such as in the event of physical death.

5.7.1 Protecting the Administration of Identity Data

An Operator must provide an individual with the ability to alter the individual's identity attributes and to configure how identity attributes flow in and out of the INDI environment. An Operator may allow for an individual to delegate part of these configuration tasks to another individual or Operator. Typical tasks that an individual may wish to delegate to the Operator are the linkage with authoritative identity attribute sources. In this case the Operator will take over the administrative details on behalf of the individual for specific interactions.

Each INDI environment must therefore enforce an individual-controlled administration policy, which states which roles, or legal and/or natural persons are allowed to perform which administrative and operational functions on the individual's INDI environment. Authorizations granted by the individual are mapped onto permissions within the administration policy. State-of-the-Art technologies for policy-based access control – e.g. XACML – could be sufficient to express adequate policies and provide a framework for a secure enforcement.

Given the demand for authorization of the individual to control his or her own identity data, INDI environments must be able to verify the INDI, when consented by the User or mandated by law, and (in some cases)the identity of the individual and his or her delegates. While the INDI environment must perform the authorization and access rights enforcement on its own, it should rely on external authentication that can be traced back to an accepted trust anchor. The mecha-

nism strength and trust level of the trust anchor determines the trustworthiness of most of the attributes, claims and assertions issued and provided by the Operator.

5.7.2 Enforcing the Issuance of a Digital Identity

Only the individual can trigger an Operator within the INDI environment to issue an INDI to the individual. Thus, each Operator must be able to verify the authenticity of a requestor's claimed identity and to link it to one of its managed INDIs. Again this requires that the individual's INDI can be traced back to an accepted trust anchor. Individuals should be able to restrict the disclosure of an issued INDI to a certain purpose or service.

Established standards such as SAML assertions can act as carriers for INDIs and relevant identity attributes. In the GINI project it has to be verified how usage restrictions can be expressed using standard mechanisms, how consistent usage models can be defined by the individual and how these can be enforced. Recent state of the art is based on three general models:

- usage restrictions are defined as properties of an identity e.g. using the purpose-of-use element of a SAML assertion;
- usage restrictions are mapped onto permissions that have been granted to the party using the INDI (see epSOS for an example);
- usage restrictions are defined as policies that are linked with the identity (see eFA and IHE WPAC for examples of this "policy push" model).

5.7.3 Enforcing Access and Usage Restrictions on Identity Data

Once the individual has handed over his or her identity attributes to a Relying Party, the INDI environment must be such that:

- only Relying Parties are provided with attributes that have been authorized by the individual;
- only that subset of the individual's identity attributes required to do business, and which are in line with the corresponding specifications set by the individual, are released to the Relying Party;
- additional protection demands for certain attributes as defined by the individual are enforced properly, e.g., age provided instead of date of birth.

Granting access to an individual's identity attributes to a Relying Party requires in the first step that the Relying Party can use an INDI to prove its own identity. This again requires that the Relying Party provides a service identity claim that can be tracked back to an accepted trust anchor.

In a second step, the Relying Party must prove the legitimacy of its request. This can be done in three ways:

- the individual has registered the Relying Party with the INDI environment as a known and trusted service. Common technologies for directory services and rights definition can be used to implement this pattern.

- the Relying Party provides the individual's INDI with the request in a such a way that the INDI operator can verify that the respective token was issued for the individual and can only be known by the Relying Party if the individual handed it over to the respective service. In general, proof-of-possession technologies as, e. g., provided by SAML assertions, can be used to implement this pattern. Recent implementations and standards profiles are very inflexible on this, as they require that an authorized party is already determined at the time of the issuance of the INDI. While GINI use cases require more flexibility, further investigation on the implementation of this pattern will have to be done.
- a request is considered to be legitimate if the usage restrictions linked to an INDI are fulfilled by the requesting party. This pattern requires for an attribute based access control that can be implemented using recent standards such as XACML.

In the GINI project compliance with use cases and possible security risks must be assessed for all three options.

In a third step the Operator determines the identity attributes required by the Relying Party. Again, different possible realizations of this have to be considered in the GINI project. In the easiest case the decision on the disclosure of a set of identity attributes is built upon a static decision base:

- The Operator maintains a registry of Relying Parties that holds information on the attributes required by the respective services.
- The INDI environment allows the individual to define on a per-attribute granularity which of the individual's identity attributes are disclosed for which purposes and/or service transactions.
- The Operator provides templates that define which attributes are typically required by certain types of services (e. g. eGovernment services, eHealth services).

More dynamic measures are based on negotiation between the acting parties:

- The Relying Party states which attributes it needs. This pattern is e.g. defined in the Shibboleth framework.
- The Operator and Relying Party do a negotiation on the attributes that are provided to a service. The individual is involved in the negotiation process and may accept or deny the disclosure of a certain attribute.

Static and dynamic measures can be combined. Based on the GINI use cases, trust models and security requirements different combinations must be assessed by the GINI project. Flexibility and ease of use are major requirements to be met.

In the final step the INDI environment determines which attribute services to contact in order to obtain the attributes that are about to be disclosed to the Relying Party. During this step requested levels of assurance and accuracy of the requested attributes must be considered.

5.7.4 Redress Mechanisms

In the INDI environment, redress mechanisms for dispute resolution evolve as part of the critical set of services offered by the Operators. It is envisaged that secure audit trails, with associated recovery mechanisms, will form part of the solutions offered by Operators, allowing individuals

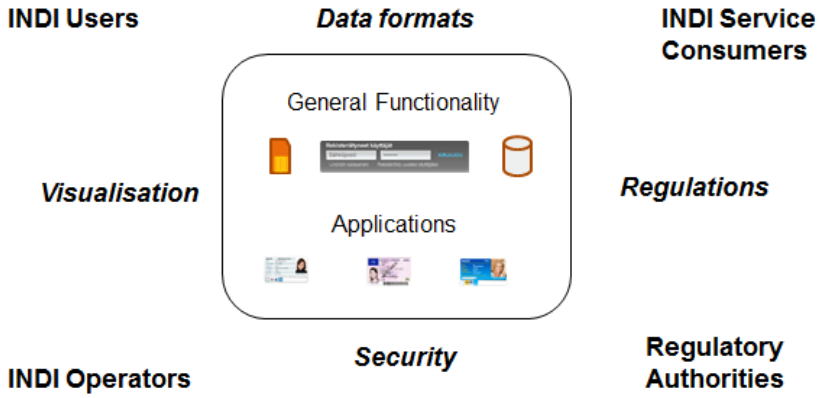
to obtain the necessary tools for exerting control of how their INDI is being used, and to be able to "observe" and "record" what happens at different stages of their interactions. On this basis, they would be able to determine their preferred levels of privacy and integrity. As a way of enabling trustworthy recovery, the Operators can be selected by individuals themselves based on their specific requirements.

5.8 Characteristics of the INDI Environment

INDI is a new infrastructure, which means that basically everything is missing. However, in order to make a gap analysis from the User point of view, at least the functional specification and market structure need to be analysed. It is obvious that there will be need for standards and regulatory framework, but they cannot be defined properly, before at least the functionality has been defined.

5.8.1 Functionality of the INDI Infrastructure

Figure 26: INDI infrastructure overview



The INDI infrastructure has several different stakeholders, who have different views of the value of the infrastructure. Because of their different motivations, they also have different views of the functionality of INDI and also how to build a good architecture. In a user-centric approach, the interests and concerns of the individual should naturally be prioritised.

First, some common functionality needs to be agreed between the parties, who will run and maintain the infrastructure. The functionality needs to be analysed from the point of view of at least the following stakeholders: users, Operators and Relying Parties.

Functionality can and should be divided into general functionality and applications.

5.8.1.1 General Functionality

There will be some general functionality, which most users should be able to utilise. Examples of general user functions are:

- Signing up to one or more Operator.
- Authenticating and signing with help of the Operator.

- Controlling the data which is visible to others (including option for anonymity).
- Storing of own data or applications.
- Showing or sending of data.
- Verifying data from others.
- Tracing data in digital transactions.

5.8.1.2 INDI Applications

In parallel to the general INDI functionality, there probably needs to be a possibility to create applications. There might be many views of what an application is and what should be regarded as general functionality, but at least two aspects can be considered, when application feature is designed:

- User view should be the most important view, because they will eventually decide, if and how they use the applications.
- There should be a possibility to define application specific rulebooks. Operators should keep the role of the definition and maintenance of the core infrastructure, but it is practically impossible to control all application specific rules. It is also possible that different applications are governed by different laws.

Applications are something, which make it possible to create and build something new on top of the INDI architecture. From the EU development point of view, there are some interesting international identity related applications, which might be implemented quite early.

Some examples of potential international applications are:

- Proof of place of residence and citizenship.
- Driver's license, student cards.
- Proof of education, proof of employer.
- Tax-related applications.
- Anonymous checkout functionality.

5.8.2 Business Models in an INDI Market

Figure 27: INDI market actors

As INDI is a new infrastructure, there is no INDI market. However, if INDI allows private organisations to become Operators, a market needs to be created. By default, the INDI market is free and it should create itself with help of free competition. However, there are two practical aspects that need to be considered:

5.8.2.1 Two-sided market

In a two-sided market, the user contracts with Operators are separated from the Relying Parties contracts with Operators. This means that the terms and conditions of the two sides of the market can also be different. Sometimes, like in card payments, only one side of the market (merchants) pay the fee and the money is transferred to the operators in the other side of market (credit card issuers) with help of transfer fees.

Authentication service markets have traditionally used so-called three-corner models. The User and Relying Party both make a contract with the Operator, who acts as a trusted third party and enables authentication of the Users. This often leads to a market where Relying Parties need to make contracts with all large Operators and because their user base does not overlap, the Operators are not true alternatives to each other. This development does not promote competition (but could still be preferred model of some operators, if they can decide).

More sophisticated models can be found from the card payment networks or GSM network. In the so-called four-corner model Operators co-operate in such a way that the Operators of the User and Relying party can be different. This makes the market more effective, because the Relying Party can always access all Users of the market regardless of which Operator it uses. Similarly, the Users can access all services with one Operator and can change Operator without losing any services. The four-corner model enables more competition between the User Operators and between the Relying Party Operators and could be the target of INDI market.

5.8.2.2 Transfer fees

In the card payment industry, the problem of the asymmetric two-sided market was solved with the transfer fee, which was defined by the payment scheme (e.g. Visa). This means that the merchant pays fee (e.g. 1%) of the payment transaction to the bank of the merchant (acquiring bank). The acquirer gives 0,7% to the scheme and keeps 0,3%. The scheme keeps 0,3% and gives 0,4% to the card issuer (the bank of the consumer). The card issuer collects the full fee and transfers money back in such a way that all parties collect their fees before money is given to the merchant.

The challenge of the transfer fee is that the scheme part (e.g. 0,7%) is easily locked and used to protect the earnings of one side of the market. The problem in card payments became so bad that the European Commission started actions against the card schemes and started to demand

an arrangement, where there are no transfer charges. This would lead to user fees for the card payments, which has proved to be very difficult in the existing market, where users do not pay.

For the GINI project, it is obvious that transfer fees should be examined with a view to finding the best possible solution for the user, while creating a competitive open market and avoiding vendor and technology lock-in.

5.8.3 Standardization considerations

It is quite obvious that some standards are also missing from INDI architecture. Traditionally, the identity related standards have defined:

- Security protocols.
- Data formats (often XML schemes).
- Control flows of an on-line identity transaction.

In the INDI environment however, the following standardisation aspects might have more relevance:

- Visualisation standards (how do user recognise the trust infrastructure).
- Data semantics (how are identity data fields interpreted).
- Rules related to on-line and off-line data handling (registers often function on-line, but it is not practical to require on-line access everywhere – this leads to questions of expiry of data).
- Application specific standards and rule books.
- Legislative standards within EU.

Precisely which standards that the Operators and Relying Parties will use in their communication, is outside the scope of this WP, but will be addressed at a later stage in the GINI project.

Regardless of the technical standards that are still needed, a bigger interoperability issue is: who decides what standards are used by the INDI Operator and the service provider in their communication with one another. This might be established bilaterally (a completely market-driven approach) or there can be a rule-making body for the ecosystem which specifies which formats the actors must comply with? (see earlier statement of rule-book being established for certain applications).

The task at hand, down the road, will be that of determining the characteristics of this entire playing field, including its governance. GINI will examine these issues further down in WP3 and in the final White Paper (WP5).

5.8.4 Different views – Different interests

Finally, a missing piece is an entity, which can influence the INDI infrastructure and which has true motivation to protect the Users' interests. The User interest is probably quite clear. New easy-to-use services, which make it possible to do things better and more securely on the Internet in such a way that the privacy is not compromised.

Other stakeholders might have different interests, which should be considered:

- Many private Operators would probably like to have a one-vendor high-price low-cost system, which locks in the Users – in a competitive environment, the Operators like to build obstacles for changing the Operator.
- Relying Parties would like to outsource all difficult parts of the identity management but they would still want to have access to Users' personal data – prices of the INDI services should be as low as possible (even free for Relying Parties) and transfer as much liability as they can to Operators and users.
- Regulatory authorities are often quite security-driven and risk-avoiding – this easily leads to requirements, which are expensive to implement and provide bad and cumbersome user experience.

One of the key INDI questions is the organisation of INDI maintenance – who would protect users but still make it possible to create and develop INDI infrastructure?

5.9 INDI Use Cases

We can see the usage of the new INDI services promoted by GINI in several ways. Broadly speaking, we can see INDI services from two sides:

a) User-centric view: Describing Use cases in specific context that shows the user experience.

b) Function-Centric View: Description of new functionality that supports several user-centric UCs.

It is foreseen that the user-centric view will be important to get a better understanding of what GINI might propose that is not happening already, but also that the functional view will be important to translate the user experience into INDI services. The two orthogonal views will together form the functional scope of GINI and the dimensions of the INDI environment.

5.9.1 User-Centric Use Cases:

5.9.1.1 Use Case 1: Online Voting / Balloting

A small city wants to build a new road around the city centre. Two possible routes have been investigated. The city wants its citizens to decide on which route to build. For cost reasons the respective balloting is done online. Each adult citizen has a single vote. Voting is anonymous.

Abuse case: Voters' identity is disclosed to unauthorized entities. That causes problems for those who have different opinion than the others.

Desired properties/ requirements:

- anonymity of voters;
- identity verification (including residence and age);
- single vote for each adult citizen.

Solution: The city sets up a balloting web portal where a citizen can select between the two routes. Citizens log in to their INDI environment. The INDI environment is linked with the registration office (acting as an authoritative identity attribute source) and issues an assertion that

just states the city of residence of its owner and a claim that the owner is over 18. The authenticity of both attributes can be verified by the consumer of the assertion. In order to prevent multiple votes by the same person, the balloting portal requests that a transaction code is included with the assertion that is unique for the citizen within the context of the current use case.

Value: User value comes from the better opportunity to participate to the decision making. Also, both the user and the authority save costs, when the balloting can be done online.

Business model considerations: There is clear value for both parties, which makes it possible to charge for both sides of the market. However, in the user side, the charging per transaction cannot probably be justified. At the same time, charging for the general use service, which could also be used for voting, can be justified.

5.9.1.2 Use case 2: Public transportation – welfare benefits

A public transportation company has an online portal where citizens can buy tickets. There are 3 prize levels: normal, children and reduced price. Students, seniors, disabled people and welfare recipients are eligible for fares at reduced price. However, this portal is rarely used by welfare recipients as they are afraid of they might be stigmatized by disclosing their social status.

Abuse case: social status is disclosed to unauthorized entities or is stored longer than necessary and is disclosed to colleagues after recipient has found job.

Desired properties/requirements:

- assurance that only eligible persons receive the benefit;
- without disclosure of welfare status beyond those entities that need-to-know;
- no (opportunity of) storage for entities that strictly do not need to store welfare status (in other words, prevent storage in non-authentic sources).

Solution: The citizen logs into his INDI environment. The INDI environment issues an anonymous assertion that states if the reduced price applies to the owner of this assertion. The ticket price is paid using normal online banking where the bank issues an anonymous confirmation that a certain amount of money has been transferred to the transportation company.

Value: User value comes from the better utilisation of the benefits together with the improved privacy. On the service provider side the solution improves privacy and service.

Business model considerations: There is value for both parties, which makes it possible to charge for both sides of the market. However, in the user side, the welfare benefits already indicate that there is limited amount of money available. Because the user group volumes are probably reasonable, this type of service might be free for users.

5.9.1.3 Use case 3: Job-related information and legal requirements

Use Case 3.1: Student Job

A student looking for a job is expected to provide information about his status (student), health insurance and nationality. While the student does have physical tokens (student card health insurance card, national id card) attesting to these various properties, he would prefer to fill in a number of applications online without having to submit photocopies. From their part, the (prospective) employers do not want to spend their time sifting through bogus applications.

Abuse case: in the Member State in question, health insurance providers have some religious or political affiliation. Disclosing his health condition through the insurance provider to an employer may affect his selection for a job.

Desired properties/requirements:

- from the student's perspective: easy management of personal information, "own" his personal information and disclose to whom he wants regardless of which service provider or governmental department is actually holding it;
- data minimization: do not disclose more than what is strictly needed for purposes of the transaction between each party;
- from the service provider's perspective: assurance of reliability of information to avoid wasting resources in the recruitment process (but there is willingness to rely upon less-than-real time information?).

Solution: A student can link both the university enrolment office and the health insurance company as authentic identity attributes sources with his INDI environment. The INDI environment can either automatically request current attestations on demand or takes care that every 3 months a new attestation is requested. The INDI environment can issue an assertion that expresses all the requested status information for employment.

Value: User value comes from the better utilisation of own verified information and increased trust in the evaluation of the User. The User has also privacy benefits, if only necessary information is shown.

Business model considerations: There is value for both parties, which makes it possible to charge for both sides of the market. As this kind of document is not used constantly, transaction fees might be justified.

Use Case 3.2: Job-related attestations cross-border

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 must submit a certified copy of his grades and his degree. He also needs to supply 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 in question to issue the necessary attribute assertions to the prospective employer.

This use case will be influenced from the SPOCS project which implements Single Point of Contact services aiming to support the implementation of the Services Directive.

Value: User value comes from the better utilisation of own verified information and increased trust in the evaluation of the user. The User has also privacy benefits, if only necessary information is shown.

Business model considerations: There is value for both parties, which makes it possible to charge for both sides of the market. As this kind of document is not used constantly, transaction fees might be justified.

5.9.1.4 Use case 4: eHealth platform

Mary is a retired woman who was in a car accident that severely injured her leg. As a result she needs to see a physician and a specialist on a regular basis. To help facilitate the recovery process as well as possible, the physician and the specialist need to be in close contact with one and other to exchange information about her condition and how she is responding to treatment. Mary wants to know what is in her health record but at the same time she does not know much about computers. She wonders whether there is some way for her to allow her daughter, Sarah to access her information on her behalf.

Abuse case: the same functionality that supports delegation is used to elicit medical information from Mary for illicit purposes …

Desired properties/requirements:

- from Mary's perspective: delegation service needs to be reliable and protect her privacy
- from Sarah's perspective: usability and integrity of process
- from the physician's and specialist's perspective: access to medical information;
- from the hospital's perspective: confidentiality and security of processes and information?

Solution: The hospital delegates access rights for Mary's medical records to Sarah's INDI. At the same time, the hospital ensures the integrity of the process by only allowing physicians and specialists involved in the process to have access rights through their INDI. This is also further ensured as the hospital has traceability throughout the process, in case of misuse of Mary's medical information.

Value: Largest value is probably tied to the automation savings in the communication between the physician and the specialist. For the User, the benefit is to have control over her own information, which often would not be accessible.

Business model considerations: This use case shows clearly the difference between a business User and a consumer type User. It is quite clear that the business Users are ready to pay more for the INDI service in this case. However, there is clearly value, which could be charged for, for all Users.

5.9.1.5 Use case 5: ePortfolio

Bill has been working for his current employer for more than 15 years. He is more or less happy in his current position, but with all his experience he would be interested to know what type of positions he might be able to obtain outside of his company. He would like to put an ePortfolio online indicating his wish to receive job inquiries, but he is worried that if his current employer gets wind of this he will be shown the door. On the other hand, the service provider that distributes the ePortfolio has a clear policy which states that all information provided must be truthful, as they want to offer their business partners (recruiting departments, head-hunters) assurance that the profiles they are viewing are reliable.

Abuse case: Bill's ePortfolio and job query is disclosed to his current employer and Bill is fired.

Desired properties/requirements:

- from Bill's perspective: reputational protection guaranteed by the service provider that distributes the ePortfolio;

- access to functionality without being directly identifiable;
- disclosure of actual identity only once there is an interest in which he himself is also interested;
- from the ePortfolio SP: ability to verify the authenticity of Bill's information or prove that Bill assured that the supplied information was truthful.

Solution: Bill uses his INDI environment for establishing his ePortfolio to receive enquires for new job offers. The INDI environment makes it possible for Bill to be anonymous, but still provide relevant information for other companies to view. Most importantly, both the service provider distributing the ePortfolio and Bill can link their INDI to trusted third party(ies), which provides traceability of what happens at different stages of their interaction. As a result, Bill can hold the service provider accountable if his information is leaked to his current employer. On the other hand, the service provider can also hold Bill accountable if the information he provided was false.

Value: In recruiting process, both sides have clear value.

Business model considerations: Also in this case, consumer type user interacts with a business user. The business users might be willing to pay more for the service or even finance it totally in this case.

5.9.1.6 Use case 6: Person-to-Person Transactions

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.

Abuse case: One of the sellers regularly sells used books about discovering homosexuality, and is later "outed" by a co-worker.

There will obviously be conflicting interests and sharply varying properties/requirements from the seller's perspective, the buyer's perspective, and from GINIbay's perspective.

Value: In person-to-person transactions, the seller has a need to proved the credibility and the value is higher. The identity of the buyer is only needed, when the commitment to buy is particularly important (value might be high).

Business model considerations: Both parties could be charged, but the buyer would probably be most willing to pay for the check of the seller.

5.9.1.7 Use case 7: Pharmaceutical Research

Once a pharmaceutical organisation identifies a molecule that has a greater than x% probability of delivering results for commercial gain, it starts a very long process of R&D before it can go to the open market.

At an early stage, to expedite the drugs development, pharmaceutical organisations introduce specialist life sciences organisations to assist. This has the very real possibility that if staff gets fired, they steal Intellectual Property (IP) belonging to the organisation employing them.

Abuse case: Pharmaceutical organisations have suffered large-scale Intellectual Property Rights (IPR) issues when the employees they fired started working for another organisation in the relevant field.

Desired properties/ requirements:

- IPR protection;
- access rights and responsibilities;
- traceability.

Solution: Employees in Pharmaceutical organisations use INDIs with specified access rights and responsibilities according to their role. In this sense, the employees are only allowed to access what is mandated by their role. If the pharmaceutical organisation fires employees, they can also immediately terminate employees' INDIs with associated access rights. To protect its IPR, the Pharmaceutical organisation can use internal trusted third parties with trustworthy recovery mechanisms to witness sensitive digital interactions.

Value: The value is clear for the pharmaceutical company, which could finance the whole solution.

Business model considerations: This is a good use case, where the company might pay the INDI fee of the user because they need the service for own protection and privacy purposes.

5.9.1.8 Use Case 8: University registration portal

Tom registers for an examination on the university's registration portal. Later, when he arrives to take his exam, the examiner tells him that his name is not registered for the examination.

Abuse case: Tom has registered for the examination but is not allowed to take the exam since the university does not find any trace of his registration.

Requirements/suggestions:

- Tom's perspective: possibility to obtain valid proof that he registered for the examination;
- university perspective: possibility to have traceability so as to verify the proof presented by the student.

Solution: Tom uses his INDI environment and registers for the university examination. A trusted third party witnesses the registration and provides a valid receipt to both Tom and the university, containing the necessary details. The trusted third party also stores a digital representation of the interaction, in case of a dispute.

Value: The greatest value is in the improved protection of the rights of the User.

Business model considerations: There is value to both parties although the value seems to be greater for the User and they might be more willing to pay.

5.9.1.9 Use Case 9: Communication Service Provisioning

Story: Communication Service Providers are by law required to keep all connection data of their customers. This is a violation of citizens' privacy as it allows the provider to profile its customers (even though this is not allowed there have been many cases where it has been done). Communi-

cation providers claim that they cannot allow for anonymous use of their services because they must be able to disclose a customer's identity in case of a request from an authorised governmental agency.

Technical Solution: A citizen's INDI environment can create a pseudonymous identity for this citizen. The INDI environment guarantees that it can disclose the pseudonym to authorized governmental agencies. The citizen can use this pseudonymous identity for communication services. In case of an investigation the communication service provider hands the pseudonym to the authorized governmental agency which can request disclose of the pseudonym from the INDI Operator

Value: The largest value is in the improved protection of privacy. For the operator, the value is probably negative.

Business model considerations: This is a good example of a case, where other party has negative value and the willingness to pay would be zero. The only solution for the communication service providers is that the functionality is required by law and there are several INDI operators in the market.

5.9.1.10 Use Case 10: Social Networking

The same as Use Case 9. A citizen can participate anonymously in a social network. In case of misbehaviour the provider of the social network can request the INDI environment to disclose the real identity.

Value: The largest value is in the improved protection of privacy. In this case, there is also value for the service, because it improves the handling of the misbehaviour.

Business model considerations: Both sides have value and could be charged.

5.9.1.11 Use case 11: Online petition against/for teaching of creationism

Story: A group "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 the every resident is able to sign the petition once.

Abuse case: profiling of religious preferences

Desired properties/requirements:

- Anonymity;
- assurance of residence;
- every resident only able to sign once.

Value: The greatest value is in the possibility to arrange this kind of petitions in a cost-effective way without compromising privacy.

Business model considerations: In this case, the arranger of the petition would probably be most willing to pay. However, there is value for the User as well.

5.9.1.12 Use Case 12: Renewal of Authoritative Documents

Story: 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.

Technical Solution: The citizen's INDI environment is linked with all the sources of authoritative documents and special-purpose smart cards (which are identity attribute sources, too). It keeps track in when the validity of a document or card ends and automatically collects all attestations that are required for renewal. If the citizen confirms, the INDI environment issues the request for renewal of a document or smart card and by this even keeps track that all information it gathers from external identity attribute sources is current and valid.

Value: There is a large value in the automation of an expensive manual process.

Business model considerations: In this case, both sides could be charged, because the value is obvious.

5.9.2 Function-Centric Use Cases: INDI user services

5.9.2.1 User Service 1: Presentation of own verified data to services or persons on the Internet

Motivation and Objective: Verified person-specific documents (e.g. identity document, proof of place of residence, etc.)

Main Requirements:

- The User has a need to prove something about him/herself:
  - Authority request (identity card, driver's license, student card).
  - In order to improve own credibility (place of residence, picture).
  - Credit rating (credit check from the source country, income/tax certificate).
- In order to be credible, at least two trusted parties are needed to guarantee the data, which is shown:
  - Trusted third party, which guarantees the integrity of the data (e.g. INDI operator).
  - Trusted source of the data (e.g. population centre of a country), which creates the credibility of the data.
- User needs a method to show the data:
  - During on-line transaction or conversation.
  - With help of off-line messaging.
  - With help of off-line links (link is user profile for example).

Main issues to solve:

- What is a driver's license (or some other data) from the User's point of view?
  - Document, application, set of attributes?
  - Do I have to pay for it and to whom?
- What does the trusted data look like to me and to the viewer?
  - What should be standardised in visualization?
- How is the trust chain built from the user perspective?
  - How can get my data in a trustworthy form?
  - Will others trust the data source, which I select?
- How can I know to whom I can present my trusted data?
  - It is a reasonable request that the receiver of the data is also trusted.
  - Basically, no sensitive data should be transferred over an untrusted network.

5.9.2.2 User Service 2: Verification of the identity data of other users (or organisations)

Motivation and Objective: Beyond the technical verification of electronic signatures

Main Requirements:

- The user or service has a need to check the data that somebody has presented:
  - Civil servant in some government process.
  - Seller, buyer or renter of some service of product.
  - Users of recruiting or dating services.
- In order to check anything, the data or some other link to the user must be received:
  - Document, link or some data.
  - Again, two trusted parties are needed to create the credibility.
- On-line checking would be nice but not necessary mandatory:
  - Data sources seem to allow use of three months old data.
  - On-line refresh should be possible, but would probably cost something.

Main issues to solve:

- Who verifies the data, which somebody else shows to the user?
  - Software, service provider, data source?
- How old data can be presented and verified?
  - Is there a need for a data format, which includes on-line refresh functionality?
- How is the trust chain visualised?

  - Are there one or more trust schemes?
  - How do I recognise "my" trust scheme?
  - Do the trust schemes inter-operate?
- Who takes the responsibility of the verification result?
  - How can responsibilities be assigned so as to ensure the integrity and validity of the data?

5.9.2.3 User Service 3: Verified identity and Internet communication

Motivation and Objective: How to trust others in messaging and chatting.

Main Requirements:

- Communication varies from on-line to off-line:
  - Voice or video call.
  - Chatting.
  - Messaging (sms, e-mail, etc.).
- In the electronic communication it is typical that there is a link to the parties at the other end (phone number, nickname, e-mail address):
  - However, the current technologies do not verify anything or enable verification.
- In addition to the verification of the parties, there might be a need to verify the communication itself:
  - Recording and signing of a conversation.

Main issues to solve:

- Can we build trusted communication on top of existing communication technologies?
  - Trust services based on addresses (phone number, e-mail address, etc.).
- What would a new "trusted communication network" look like to the User?
  - Let's make the deal in the GINI network?
  - Can you cc that to my GINI account?
  - Could you send me your GINI card?
- Are the INDI operators like telecom/Internet operators or are they forming a new "trust industry"
  - Is there a regulation framework, which Users can understand?

5.9.2.4 User Service 4: Creation and use of a verified anonymous identity on the Internet

Motivation and Objective: Anonymous (but verifiable internet transactions, e.g.: Shopping - Basic right for any Internet user?

Main Requirements:

- Current practises of Internet registration do not respect privacy very much:
  - Often it is necessary to give data, which is probably not used or needed but which is collected for marketing or other purposes.
- Many Internet users already register using totally false data because they are annoyed about the requests:
  - All services should include an option to use the service anonymously.
- How can you be anonymous, but still trustworthy for a merchant or a service provider?
  - Anonymous identity cards.
- It should be a basic right in Europe to be anonymous on the Internet.

Main issues to solve:

- How can you be anonymous but trustworthy in practise?
  - Somebody must know you to build the trust.
  - It should be possible to verify an anonymous identity.
- How do the merchants and service providers see the INDI network?
  - Is an INDI address enough to create a user entry even when the User does not reveal any data?
- How do you switch from anonymous to non-anonymous?
  - Restricted to person-to-person or person-to-service event.
  - Reasonable request is that both reveal their anonymity unless there is a good reason, why the other stays anonymous.

6 Abbreviations

epSOS	European Patients Smart Open Services (FP7 Large Scale Pilot)
GINI	Global Identity Network of Individuals
HCP	Healthcare Professional
ICT	Information and Communication Technology
IDM	Identity Management
IdP	↑Identity Provider
INDI	↑Individual Digital Identity
PEP	Policy Enforcement Point (→XACML)
PIP	Policy Information Point (→XACML)
PAP	Policy Administration Point (→XACML)
PDP	Policy Decision Point (→XACML)
SAML	Security Assertion Markup Language (→SAML)
SOA	Service-Oriented Architecture
SSO	Single Sign-On
STORK	Secure Identity Across Borders Linked (FP7 Large Scale Pilot)

Annex 1: Glossary

Introduction

This glossary is produced to ensure consistency of language, terms and definitions in all documents of GINI.

As much as possible, new definitions have been developed, tailored for GINI. The remaining terms or definitions have been taken primarily from established international standards, wherever possible.

Terms and definitions

Term	Definition
Access right(s)	Permission given by a legal person to the owner of an INDI for accessing their systems.
Accountability	In its core meaning, implies giving an account of performance to someone entitled to demand that account. It includes providing information (transparency) as well as accepting remedies and sanctions in the case of unsatisfactory performance.
Accreditation	The action granting an authority to perform a defined service.
Accreditation Authority	Assesses and validates that identity providers, attribute providers, relying parties, and identity media adhere to an agreed upon Trust Framework.
Anonymous	Not named or identified. Anonymous transactions allow for information exchange between parties without the need to identify the parties involved.
Attribute	A named quality or characteristic inherent in or ascribed to someone or something. Attributes can include personal qualities (e.g. age), ambient information such as location, or certifications that serve as proof of a given capability.
Attribute Provider	Responsible for all the processes associated with establishing and maintaining a subject's identity attributes; they provide assertions of the attributes to the individuals, other providers, and relying parties.
Audit	Systematic, independent and documented process for obtaining audit evidence and evaluating it objectively to determine the extent to which the audit criteria are fulfilled. Independent review of the system and its operation to assess compliance.
Audit trail	The chronological set of records that provides evidence of system activity. These records can be used to reconstruct, review and examine transactions from inception to output of final result.

	NOTE: The list can be generated by a computer system (for computer system transactions) or manually (usually for manual activities).
Authentication	The process of verifying a claimed identity.
Authorization	The process of ensuring that the user of a network has received permission to use the facilities of the network that are in his or her access profile. Authorization follows the processes of identification and authentication.
Availability	The capability of a system to ensure that the required information is available whenever required and subsequently that it is able to perform its assigned tasks within an acceptable amount of time.
Best practice	A technique or methodology that, through experience and research, has proven to reliably lead to a desired result.
Biometrics	Automated methods of recognizing an individual based on measurable biological (anatomical and physiological) and behavioural characteristics.
Brokered identity	Relies on trusted third parties (brokers) acting on behalf of the identity owner.
Brokered trust	One party implicitly trusts the other partner despite having no direct trust relationship to each other, by the mediation of one or more intermediaries.
Certification	Procedure by which a Certification Body gives assurance that all or part of products, processes, systems or persons conforms to a set of requirements.
Circle of Trust	A trusted group of identity and service providers who share linked identities and have pertinent agreements in place regarding how to do business and interact with identity providers.
Claim	An assertion of the truth of something, typically one which is disputed or in doubt (e.g., a digital identity or other attribute).
Claimed identity	The digital identity claimed by the owner of an INDI in a digital interaction.
Community trust	Two parties create a valid trust path by their enrolment in a certain authentication community and a subsequent acceptance of its norms and practices. Apart from the communities established authentication norms and means, no additional intermediaries are introduced or used.
Confidentiality	Preserving authorized restrictions on information access and disclosure to prevent disclosure to unauthorized individuals, entities or processes, including means for protecting personal privacy and proprietary information. Consider following alternative:

	security objective is understood as keeping the content of information secret from all entities except those that are authorized to access it. (see Menezes, A.J., Van Oorschot, P.C. and Vanstone, S.A., Handbook of Applied Cryptography, CRC Press, Boca Raton, 1997, 32)
Cybercrime	The use of computers, computer systems, hardware devices, networks and/or the internet to perpetrate fraud and other crimes against users.
Data curation	The actions needed to maintain research data from point of creation to ensure they are fit for contemporary purpose and available for discovery and re-use. Implicit to this are the processes of archiving and preservation. Higher levels of curation will involve maintaining links between datasets, annotation, published materials and other information resources.
Decoupling	Two or more systems that are able to transact without being connected, or coupled.
Delegated Role	A role assigned to an individual user with that user's agreement and that has certain rights and responsibilities relating to the delegated role.
Digital Identity	A set of attributes that characterize an entity within a particular context or domain of applicability
Device	A physical construct, generally electronic, that is capable of storing and processing information, e.g., a Personal Computer, web server, mobile phone, or smart card.
Digital information	Any information that is represented in a digital form.
Digital interaction	An interaction carried out using digital communication means.
Digital signature	A digital signature is data appended to, or a cryptographic transformation of, a data unit that allows a recipient of the data unit to prove the origin and integrity of the data unit.
Direct brokered trust relationship	When one party trusts a second party, who, in turn, trusts or vouches for, a third party.
Direct trust	When a relying party accepts as true all (or some subset of) the claims sent by the requestor.
Dispute resolution	The process of resolving disputes between at least two parties. Methods of dispute resolution include: arbitration; conciliation; litigation; mediation.
End-user	The ultimate consumer of a finished product.
Federation	The process of establishing a trust relationship between two or

	more entities or an association of entities compromising any number of service providers, identity providers or other entities.
Federated identity	Based on a conceptual separation between service providers (SP) and identity providers (IdP) and concerns the arrangements that are made among several organisations and individuals, that let entities use the same sets of identification data, to get access (and authorisation) to the several different (otherwise autonomous) services offered by all the organisations associated with the system of federation.
Global identity	Serves to identify entities in a broader context, i.e. across local domains or within one global computing ICT infrastructure, e.g. the Internet, the web or a grid structure.
Identification	The association between a person (subject) and the full name.
Identity Provider	A kind of service provider that creates, maintains, and manages identity information for principals and provides principal authentication to other service providers
Indirect trust	The affected parties solely rely on claims asserted by a common third party with which a pre-existing trust relationship is already established.
Individual Digital Identity (INDI)	An Individual Digital identity is the digital representation of a set of claims made by one digital subject about itself or another digital subject.
Individual user	A physical person using ICT who can be uniquely identified.
INDI ecosystem	Brings together institutional, market and societal stakeholders to define a legal, regulatory and operational framework capable of allowing for the controlled and appropriately regulated establishment of INDI operators in the EU.
INDI Operator Model	INDI Operator Model include provisions for: An architecture for a service-based user-centric identity ecosystem Ways to manage interoperability and data governance Traceability, recovery, auditing, accountability Legal and Regulatory provisions, including possibilities under current law and possible gaps to be addressed Requirements for Service Level Agreements with consumers
INDI environment	An environment, created by an INDI Operator for using an INDI, through which the user can manage disclosure of his/her identity and other information from one domain to another.
INDI operator	Manages and enables the INDI environment together with other INDI Operators. It acts as a gateway to the INDI environment on behalf of one or more INDI users, registers or relying parties.
Information and Communication Technology (ICT)	ICT refers to a diverse set of technological tools and resources used to communicate, and to create, disseminate, store, and manage information. These technologies, for example, include computers, the Internet, broadcasting technologies (radio and televi-

	sion), and telephony.
Information security	Preservation of confidentiality, integrity and availability of information; in addition other properties, such as authenticity, accountability, non-repudiation, and reliability can also be involved.
Infrastructure	Consists of the integrated technical components (e.g., hardware, software, networks, applications and protocols) required to deliver online services in accordance with the trust framework and the programs necessary to support them.
Integrity	A quality which implies that the items of interest (facts, data, attributes etc.) have not been subject to manipulation by unauthorized entities.
Intellectual Property	Intellectual property (IP) refers to creations of the mind: inventions, literary and artistic works, and symbols, names, images, and designs used in commerce. Intellectual property is divided into two categories: Industrial property, which includes inventions (patents), trademarks, industrial designs, and geographic indications of source; and Copyright, which includes literary and artistic works such as novels, poems and plays, films, musical works, artistic works such as drawings, paintings, photographs and sculptures, and architectural designs.
Interacting party	One party in an interaction.
Interaction	Reciprocal action or influence. NOTE: Examples of interactions are: - Financial Transaction between two parties, - Multiple parties signing a document; Non Repudiation of receipt or sending a document: Authorised access to a resource, etc.
Interoperability	The ability of disparate and diverse organisations to interact towards mutually beneficial and agreed common goals, involving the sharing of information and knowledge between the organisations via the business processes they support, by means of the exchange of data between their respective information and communication technology (ICT) systems.
Intellectual Property Rights (IPR)	Rights (formally) conferred on an individual or on a legal entity regarding the ownership and use of intellectual property.
Local identity	A digital identity that is created and used only in a closed environment or domain. A typical example is a local password-based access, where user accounts, associated groups and passwords are stored within a file in the host environment.
Jurisdiction	The official power to make legal decisions and judgements. Can also refer to the territory or sphere of activity over which the legal authority of a court or other institution extends.

Legal interoperability	The ability to carry out digital transactions in a legally enforceable way regardless of jurisdiction. It allows parties to verify identities, roles and legal status of other parties in cross-border interactions.
Legal person	Any entity which is recognized as having its own 'legal persona', i.e. a separate legal existence (e.g., a natural person, a corporation recognized as having a separate legal persona within its jurisdiction etc). Note: legal personality implies that the entity can be sued in its own right. The legal person is a legal fiction through which the law allows a group of natural persons to act as if they were a single composite individual for certain purposes, or in some jurisdictions, for a single person to have a separate legal personality other than their own.
Legal rights	All rights recognised by the law as having legal existence and effect.
Level of Assurance	The degree of confidence that the individual who uses the credential is, in fact, the individual to whom the credential was issued.
Network identity	Context-sensitive identity, attributes, rights, and entitlements, all maintained within a policy-based trusted network framework.
Non-digital world	Our existence as human beings without the use of ICT.
Non-Repudiation	Addresses the capacity of a given system or entity to ensure that the actual execution of a given event may not be disputed in retrospect.
Online	The state associated with the ability to connect and communicate with other networks, systems, computers, subjects or components in real time through the Internet.
Organisational interoperability	The coordination of processes by which different organisations such as different public administrations collaborate to achieve their mutually beneficial, mutually agreed eGovernment service-related goals.
Partial identities	A set of attributes that sufficiently identify an entity within a set of entities.
Policy	Overall intention and direction as formally expressed by management.
Privacy	The right of an individual to control or influence what information related to them may be collected and stored and to whom that information may be disclosed.
Private Policy	A set of rules/policies governing the access to an INDI. These rules are known only to the owner of the INDI and are not disclosed to anyone else, unless the owner wishes to disclose them. A private policy can be any type of policy that an individual uses to manage the INDI, e.g. security policy, back-up policy etc. The individual user can chose different private polices for different interactions.

Pseudonymization	The replacement of all data that identifies a person with an artificial identifier
Reference	The confirmation by one party that an INDI, and therefore a claimed identity, has interacted with it at some point in time. References can be obtained from anyone whom the owner has claimed that they have interacted with.
Register	An entity which maintains information about one or more INDI users. The register maintains this information for its own business purpose(s) (in the case of private entities) or public mission (in the case of governmental entities) which exist(s) independently of the INDI space.
Relying Party	Entity that relies on the veracity of a claim. Within the ecosystem, a relying party is responsible for interacting with credential, identity, and attribute providers as needed to verify parties with whom they exchange information. Connected to the INDI space through one or more INDI Operators.
Repository	A place where or receptacle in which things are or may be stored.
Role	A set of permissions granted to a user and applied to specified groups of resources.
Service provider	A company, organisation, etc. which provides a service to customers.
Single Sign-On (SSO)	The ability for a user to authenticate once to a single authentication authority and then access other protected re-sources without re-authenticating.
Standard	Documented agreements containing technical specifications or other precise information to be consistently used as rules, guidelines or definitions or characteristics to ensure that materials, products, processes and services are fit for their purpose.
Surveillance	The unwanted observation of individuals but also of communities and populations.
Third party	Natural or Legal person that is recognised as being independent of the parties involved, as concerns the issue in question.
Traceability	The ability to determine who did what and at what point in time.
Transaction	The act of transacting within or between individuals and/or groups (as carrying on commercial activities).
Trust	The subjective state of reliance of an individual human being in another individual, system or transaction that it will meet with his or her specification or will otherwise behave as expected.
Trust level	The level of trust that one interacting party has with another interacting party.
Trusted Third Party (TTP)	Security authority, or its agent, trusted by other entities with respect to security related activities.

	NOTE: The TTP is an entity which facilitates interactions between two or more parties who all trust the third party; they use this trust to secure their own interactions.
Trusted time stamp	Time stamping of information performed by, or with the direct involvement of, a party underwriting the accuracy and consistency of that time stamp. NOTE: Trusted time stamps are used to enhance the integrity or authenticity of information to which it relates. It is often used in conjunction with cryptographic digital signatures.
Unlinkability	The condition in which a third party cannot determine whether two actions or two data items belong to a single user.
User	An individual who uses an INDI to access the INDI Space and present information about herself towards other parties. The user can act in various roles, e.g. citizen, employee etc. with different rights and responsibilities. A user may also use her INDI in order to receive data about other INDI users.
Virtual Individual Digital Identity Folder (VINDIF)	One-stop interface for linking, storing, managing identity attributes and related application data/documents.
Verifier	Entity that corroborates a claim with a specified or understood level of confidence