Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática
2010

Frank Alessander de Oliveira Pimenta

IDENTIDADE DIGITAL FEDERADA GLOBALiD

GLOBALiD FEDERATED DIGITAL IDENTITY

"Strive not to be successful but rather to be of value."

⌐Albert Einstein.

Universidade de Aveiro Departamento de Electrónica, Telecomunicações e Informática
2010

Frank Alessander de Oliveira Pimenta

IDENTIDADE DIGITAL FEDERADA GLOBALiD

GLOBALiD FEDERATED DIGITAL IDENTITY

dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica do Professor Doutor Joaquim Sousa Pinto, Professor Auxiliar do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro e do Doutor Cláudio Teixeira, investigador post-doc da Universidade de Aveiro.

I would like to dedicate this work to my parents and thank them for the endless help and support.
I owe it all to them!

o júri / the jury

presidente / president	Prof. Dr. José Luís Guimarães Oliveira professor associado do Departamento de Electrónica, Telecomunicações e Informática da Universidade de Aveiro
vogais / examiners committee	Prof. Dr. Fernando Joaquim Lopes Moreira professor associado do Departamento de Inovação, Ciência e Tecnologia da Universidade Portucalense
	Prof. Dr. Joaquim Manuel Henriques de Sousa Pinto professor Auxiliar da Universidade de Aveiro
	Prof. Dr. Cláudio Jorge Vieira Teixeira investigador post-doc da Universidade de Aveiro

agradecimentos / acknowledgements

I would like to deeply thank to my Master’s advisors Professor Joaquim Sousa Pinto, PhD, and Cláudio Teixeira, PhD, for their dedication, interest and professionalism in my work. Especially, I would like to thank to Cláudio Teixeira, PhD, for his continuous patient, support and help in all the obstacles that had risen during the working progress.
I really appreciated all the help and effort they both gave me along the way.
Thank both of You!

I am deeply thankfully to my parents for the tremendous effort they have been making since the very beginning for me being capable of successfully finishing my academic studies. My graduation was only possible thanks to their help and support.
To them I am forever thankful.

Special thanks to my older brother for passing the profound interest for knowledge on me.

My dear sister, thank you for helping me in the bad moments.

My dear little niece Julia, keep talking and smiling as much as you use to, that’s living your life amazing the people that love you.

My dear little nephew Samuel, try being a successful football player first rather than an engineer, instead you will have fruitful life. And thereby you will have a higher probability of one day having the yellow airplane you want so much. Trust to what you uncle says, because he is telling you the holy truth.

Friends! Thanks to all of you for your friendship during my academic life. Thanks to you I have lived great moments in my life.

Nikki, you have been amazing me every single time.

palavras-chave

GlobaliD, Citizen Card, Digital, Identity, Management, Federation, Federated, Provider, Service, Authentication, Versatile, Anonymity, Privacy, Veracity, Trustworthyness, Accountability

resumo

O presente texto propõe uma solução para a gestão de identidade digital online tendo em conta a versatilidade, o anonimato, a privacidade, a veracidade, a credibilidade e a responsabilidade do utilizador, recorrendo para isso ao uso do Cartão de Cidadão Electrónico Nacional Português e a outros meios de autenticação públicos usados diariamente pelos utilizadores. A dissertação é composta pela apresentação do conceito de identidade e das suas particularidades, por uma análise aos vários problemas da gestão da informação pessoal online, uma análise aos vários modelos, mecanismos e especificações existentes para gerir a identidade digital online (gestão de identidade digital). Uma solução de gestão de identidade digital baseada no modelo de identidade federada e associada ao Cartão do Cidadão Electrónico Nacional Português é apresentada, descrita, analisada, avaliada e comparada com outras soluções existentes.

Por fim um protótipo de um provedor de identidades digitais federadas baseado na solução de gestão de identidade digital proposta é apresentado.

keywords

GlobaliD, Citizen Card, Digital, Identity, Management, Federation, Federated, Provider, Service, Authentication, Versatile, Anonymity, Privacy, Veracity, Trustworthyness, Accountability

abstract

The following text provides a solution for the digital identity management on the Web regarding the users' versatility, anonymity, privacy, veracity, trustworthiness and accountability by using the Portuguese National Electronic Citizen Identity Card and other publicly available authentication mechanisms users use daily. The dissertation consists of the presentation of the concept of identity and its particularities, an analysis to the several problems of managing personal information online, and an analysis to the several existing models, mechanisms and specifications for the management of the digital identity online (digital identity management). A solution for digital identity management based on the federated identity model and associated to the Portuguese National Electronic Citizen Identity Card is introduced, described, analyzed, evaluated and compared to other several existing solutions. Last, a prototype of a federated digital identity provider based on the purposed solution for digital identity management is presented.

Table of Contents

List of Acronyms

3G	Third Generation
CA	Certification Authority
eID	Electronic Identity
GSM	Group Special Mobile
HTTP	Hyper Text Transfer Protoco
HTTPS	Hyper Text Transfer Protocol Secure
ID	Identity
IdP	Identity Provider
IETF	Internet Engineering Task Force
IP	Identity Provider
PIN	Personal Identification Number
PKI	Public Key Infrastructure
SAML	Secure Access Markup Language
SP	Service Provider
SSL	Securey Socket Layer
TCP	Transmission Control Protocol
TLS	Transport Layer System
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
WWW	World Wide Web

List of Figures

List of Tables

1 Introduction

Nowadays people use internet every day, to read news, watch videos, talk to friends, upload photos to a share-gallery, update twitter status, etc. Personal information sharing is one of the most common online activities that everyone does very often, either with web applications or other users. Many times users feel forced to give up about some privacy in order to share a piece of their personal information with others. Very often they tend to reproduce the provisioning of their personal information or the information of others across many online platforms giving up about either the secure storage concerns on the information shared or in its unauthorized disclosure.

1.1 Motivation

The Web offers a myriad of services to registered users. Registering on those services requires filling a registration form, choosing a username or providing an email address and some personal data, such as name, age, or any other relevant personal information. In the end of the registration process a digital identity is created in the system. When a user registers in several websites, he/she has several digital identities (assessment credentials) to manage. With the boost of Web 2.0 and electronic business, web applications started to request increasingly more personal information to allow users to use the services they provide, either to create a social network or to buy products in any e-business. This growth of personal information exchanging between users and services forced users to have their personal information scattered across many systems online. Hence, a better management of the users' information and far more secure and reliable channels of communications are required [1]. The traditional exchanging means that users have been using to share their personal information lacks the mechanisms to provide full control over the disclosure and dissemination of the personal information they share as well as its scrutiny and displacement monitoring. The security of the exchanging means and the protection of the information integrity being exchanged are not assured either. Most of the times encrypted communications channels are not used for the transmission of the information e.g., the use of TLS. The protection of the integrity of the information being exchanged in non-encrypted channels as well as its non-unauthorized disclosure is, thus, not undertaken at all. Neither the users' privacy-safety nor their anonymity when exchanging their personal information is guaranteed. Moreover, most often there are not reliable mechanisms to certify that the information users provide to represent them is either accurate or asserted as either true or false. Therefore, more secure and efficient authentication mechanisms are mandatory in order to project the next generation of digital identity management [2-3].

This masters' dissertation reflects on the certification, integrity assurance and privacy-safety of the users' personal information, and its scattering across Web applications as well as in the users' anonymity protection by taking an approach to federated identity management concept. Consequently, it also reflects on the users' accountability online.

1.2 Objectives

This Master's thesis will study and analyze the digital identity management discipline in order to gather the most valuable features and specifications for developing an identity framework. A federated identity initiative will be used as well as several strong identification mechanisms publicly available, such as the Portuguese National Electronic Citizen Identity Card [4]. The proposed framework will be named GlobaliD. The identity framework goal is to offer an improved model for online digital identity management and to test whether the current digital identity management can be enhanced in order to give users better experience while managing and sharing their personal information and accessing the myriad of services offered online. The ultimate aim is to take a step further in the digital identity management and therefore in the privacy and anonymity safety of the online users by making it more versatile, responsible, reliable, authentic, accountable, trustworthy, integral and privacy safe, anonym and thus more secure. An identity provider prototype based on the GlobaliD framework will be presented in order to illustrate the framework aims and achievements.

1.3 Methodology

To achieve the thesis' objectives, the work will be divided into three main parts: Identity, Digital Identity Management and a proposal of a Digital Identity Management framework. The Identity concept was deeply implicated in this work thus I realized that firstly I would have to approach its notion in order to have a rich background about the subject such as: gathering the main identity characteristics, requirements and concerns, before proceed to the study of the identity management discipline itself, the subject that follows. After the identity subject was studied the management of identity would be better understood and worked out. On the approach to digital identity management discipline it will firstly be taken under consideration the problems of digital identity, and its particularities. Then the technologies and the models involved on the managing of digital identities. The study of the state of the art on digital identity management is introduced afterwards. The several initiatives for managing the users' digital identities were analyzed and explained. With the digital identity management discipline understood and the main existing initiatives addressing this subject known, a proposal for a digital identity management framework arose and an identity provider prototype to evaluate the proposed identity framework aims was developed. Finally conclusions on this work and projections about future developments of the identity framework proposed were outlined.

1.4 Document Structure

Federated identity management is a model of the digital identity management discipline. For a better understanding of the concept of digital identity management it is important that the notion of identity is well understood and comprehended in its plenitude in the first place. Therefore, the following section approaches the notion of identity in several different scenarios and perspectives and deeply analyzes its particularities as well as indicates the several definitions implied. On section 3 it will be introduced the digital identity management discipline. The concepts and the actors involved will be indicated and described. Several problems that the digital identities face in the identity management systems will be discussed. It will be approached the authentication procedure and its several parts. Technologies involved in digital identity management will be mentioned and explained, as well. Section 4 introduces two models for digital identity management. The benefits and pitfalls of each model will be analyzed. It will be also presented several techniques for managing the users' credentials information online. Sections 5 and 6 are dedicated to the federated identity matter. On section 5 it will be presented a series of requirements involved in the implementation of federated identity scenarios. The specifics of federated identity will be described. Section 6 is dedicated to a number of federated identity initiatives. Every initiative's specification will be introduced and described.

Section 7 presents the essentials of the SAML 2.0 according to its use in the development of the previously referred identity provider framework GlobaliD. This framework for implementing digital identity management will be introduced, described and analyzed in the section 8. On the section 9 it will be presented the characteristics of the identity provider prototype developed. Finally, on the section 10 it will be discussed the work made and projections about a future work on the GlobaliD framework will be presented.

2 Identity

We use our identity each day of our lives. It says about our essence, about what we are and reveals our behavioral nature. Our identity is expressed within a particular environment, based on the perception we want to address to the people we interact with and on our desired goals. Therefore the way our identity is transmitted is of our concern because it establishes the proper setting, appearance and manner in the interaction with other identities as it also shapes our image to others as well. A collection of impressions is what express our identity. They shape our performance on self-presentation and understanding of the self to others. The impressions of an identity are actions or signs made by us, provoking sensations on others in order to create certain reaction on them. It gives us an outwards appearance that can be used to convince other identities in our favor. Most of the times, whether or not we are aware of the impressions we transmit, a persona is imputed to us by others. This happens regardless of our state of mind, our lack of faith or even ignorance of our performance. Impressions are hence, the elements used to express our identity, revealing themselves as a key factor to the identity presentation, expression and interaction. Therefore it is important we can control the impressions we transmit about ourselves and the way they are transmitted to whomever, whenever and whatever our reasons are [5].

The way impressions are transmitted and controlled falls into the impression management field [6]. Impression management refers to the process through which people try to control the perceptions that other people form of them. In the virtual world, the Internet, we usually interact with people as we do in the real world. Even though it happens through the Internet communications framework and not in our real environment, the way we express ourselves is still by transmitting impressions. They are still telling how we are, or even better, how we want others to think we are. Therefore, we can say the difference between physical and digital social settings is intimately connected with the different strategies for representation of self and the kinds of social discernment they afford. In a computer-mediated world the expression of identity occurs primarily not through direct experience of the body but within the constraints of digital representation constructed by interactive systems. Hence, in this environment people have to create new ways for representing themselves and new ways of reading the signals communicated by others in order to compensate the loss of physical presence. It is important that the impressions of our Identity are according with our goals in order not to be misunderstood because they are what tell whether a social engagement with us is appropriated or not. If it is, they also tell how this engagement is better to be accomplished. Thus, it is important that the impressions about our identity are transmitted clearly, leaving no margin to avoidable misunderstandings [7].

The Internet is an infrastructure of both hardware and software that provides connectivity to computers. It's the base of the Web, a collection of inter-connected documents, named websites, linked under the rules of the HTTP protocol, used by millions of people nowadays. The Web made possible for humans of any part of the world to communicate among each other more easily and cheaply, than never before, without being restricted by their current location. Everyday individuals use the Internet or the Web for communicating with other individuals, either because social or professional reasons, by posting some sort of media type on their blog or social network or even by either video or text chatting. Thus, the Web became a very used mean by which users transmit their impressions to others and express their identity. This was possible due to a evolution from a simple and static distributed hypertext service of pages for publication of contents [8] to a complex and dynamic system of hypermedia applications [9], more user centered and adapted to the different requirements of its users and participants. The advent of Web 2.0 demanded that new rules and mechanisms should be created to manage the different parts and services of the Web. Different users have different needs and objectives. The same can be said about Web application. Therewithal other features had to be added with the user customization in mind. In order to achieve the customization of the various web services working online according to the users' demands and preferences, the concept of digital identity was introduced to represent the user in a computer-mediated environment.

To better understand the concept of digital identity, firstly the concept of identity must be understood due to its ambiguity. Identity is a deep and wide concept, with plenty of particularities, which are confusing and may even be contradictory, that when not very well understood, identity may seem ambiguous. The following subsections will provide an overview of the different parts of Identity (where digital identity is one of them) for an easy and clear elucidation. Further it will be introduced some key aspects that should be considered when projecting the next generation of digital identity and the technologies made to certify an Identity. Last, as an introduction to the next section digital identity management, the laws of the digital identity by Kim Cameron will be presented for providing a baseline of the digital identity management requirements.

2.1 Definitions

Identity is a deep and wide concept, outlined by a plurality of particularities. Therefore a myriad of definitions exist to specify each particularity that makes Identity a unit. The next subsection addresses several of definitions presented in [10-11].

Subject – within the identity context a subject may be a person, an organization, a company, a club, a corporation, software or other identity (I use more entity than subject regarding this master's thesis).

Attribute – self inherent feature of an identity. An attribute of an Identity is an element that characterizes the Identity. It can be an extrinsic or intrinsic characteristic of the Identity. Medical history, past purchasing behavior, bank balance, credit rating, age, weight, and height are a few examples of any Individual's Identity characteristics.

Identifier – any attribute that solely represents the identity. The identifier of an identity distinguishes it within the context of a specific namespace. An identifier can be also referred to as name, labeler or designator.

Persistent Identifier – a persistent identifier associated with a characteristic of a subject that is difficult or impossible to change. Date of birth and genetic code are two examples of persistent identifiers related to a human being. In such case, the persistent identifier may be called as personal identifier, as well.

Transient Identifier – or one-time Identifier, is a type of opaque identifier that is only valid during a specific session. This allows the user not to be recognized at multiple uncorrelated visits to the same service provider. Since the used identifiers are different at each user visit, this identifier cannot be used by the service provider to correlate the user visits or to discern anyone of them.

Preference – is a demonstration of distinction about a particular element by the Identity. A preference of an identity can be the preferred seating on an airplane, currency used, favorite brand of shoes, and the use of one encryption standard over another, and so on.

Trait – is a self-inherent attribute that outlines the individual, such as blue eyes (being), how and where a company is incorporated (location), color of the skin, and so on.

Claim – is an allegation a subject does about a self-attribute that is in doubt or being in dispute.

Identification – is an association of a particular identifier with an individual presenting certain attributes. It may be called clamming proofing, as well.

2.2 Particularities of the Identity

Identity may be seen by different perspectives, either philosophical or mathematical. A bad understanding of the identity definition may, in certain scenarios, blur the identification of different identities. Next, some of the perspectives to which Identity may be projected will be introduced and described, some scenarios will be presented and problems that may blur the identification of different identities will be pointed out.

2.2.1 Identity

The meaning of Identity seems to be pretty clear to everyone but unfortunately it is not always the case. Any dictionary states that the meaning of the word Identity is: the Circumstance of an individual being the one that he/she says he/she is or the one that another identity assumes he/she is. If we pick the word ”circumstance” from the meaning of Identity we realized that circumstance is all that surrounds an element, all that makes part of it and belongs to it, all that somehow is connected and related to it and makes it an unique entity. Identity can be therefore briefly defined by the set of permanent or long-lived temporal attributes associated with an entity. Identity is usually what represents an entity within a specific context. The entities attributes are what identifies the entities, describes its characteristics and differentiates them from other identities. If any of the attributes of an identity is unique, it can be seen as an identifier, and therefore may be used to refer to the identity. It is important to deeply understand the Identity concept because it is what represents us as an entity. Identity is therefore important. It is related to the existence of objects, their uniqueness and distinctness from other objects. It is the relation that states the sameness or identicalness of an object and it is the fundament for reasoning and understanding. Identity is the fundament for the human being comprehension of nature and interaction among themselves. Identity is what allows individuals to position themselves and to define the relations with other objects in the environment they are within [12].

Identity may be differentiated in quality identity and quantity identity. Qualitative identity highlights the properties of the object, so objects can be more or less qualitatively identical. Numerical identity requires the absolute qualitative identity [13]. Although a Ferrari and a Lamborghini are qualitatively identical because both share the property of being a car, two Ferraris can be more identical because both share the fact of being of a Ferrari brand.

"x and y are to be properly counted as one just in case they are numerically identical" (Geach 1973).

2.2.2 Logical Identity

Identity is usually intuitive as it begins when one recognizes other. Identity was first formalized by:

- Aristotle's Law of Identity: A is A for any A. Everything is itself.

Along with two others laws it constitutes the foundation of formal logic.

- The Law of contradiction: Not (A and Not A). Nothing can both be and not to be.
- The law of Excluded Middle: A or Not A. Everything must be or not to be

The combination of this three deduces that:

- A is Not A - for any A.

Discrete Mathematics defines Identity as a relation of equivalency, i.e., for two elements to be identical they must be equivalent [14]. Equivalency is a relation of reflectivity, symmetry and transitivity:

- Reflectivity - for every A, A $\Leftrightarrow$ A
- Symmetry - for every A and B, if and only if A $\Leftrightarrow$ B then B $\Leftrightarrow$ A

- Transitivity - for every A, B and C, if A ⇔B and B ⇔ C then A ⇔ C.

Although the laws and definition mentioned before are quite intuitive and useful for recognizing identities, they are not rigorous enough for systematic recognition. Therefore more formal laws are required. The Leibniz's law, a.k.a., Identity of Indiscernibles states that:

- No two distinct substances exactly resemble each other.

This can be understood as:

- No two objects have exactly the same properties.

The Identity of the Indiscernibles is outlined by two principles: the indiscernibility of Identicals and the Identicals of Indiscernibles [12-13].

Principle 1: The indiscernibility of identical

For any x and y, if x is identical to y, then x and y have all the same properties:

$$\forall x \forall y\, [\, x = y \longrightarrow \forall P(Px \leftrightarrow Py) \rightarrow x \neq y]$$

For any x and y, if x and y differ with respect to some property, then x is non-identical to y.

$$\forall x \forall y\, [\neg \forall P(Px \leftrightarrow Py) \rightarrow x \neq y]$$

The indiscernibility of identicals states that if two objects are numerically identical (the same one), they must have the same properties, i.e. qualitatively identical. Numerical identity must imply qualitative identity.

Principle 2: The Identity of Indiscernibles

For any x and y, if x and y have all the same properties, then x is identical to y.

$$\forall x \forall y\, [\forall P(Px \leftrightarrow Py) \rightarrow x = y]$$

For any x and y, if x is non-identical to y, then x and y differ with respect to some property.

$$\forall x \forall y\, [x \neq y \neg \forall P(Px \leftrightarrow Py)]$$

The identity of indiscernibles says that if two objects have all the same properties, i.e. qualitatively identical, they are numerically identical. Qualitative identity implies numerical identity. Therefore, from the combination of the two principles it may be realized that numerical identity is equivalent to qualitative identity.

Although the Law of Identity seems to be simple and reasonable, it is philosophical controversial due to its paradoxes. I will refer just The Paradox of Time and Change, The Infinity Problem, The Ship of Theseus Paradox. It is important to take a look to these paradoxes because they give us a better understanding of identities and how to distinguish them.

The **Paradox of Time and Change** takes the human oldness to state that the Indiscernibility of Identicals cannot be used to determine whether it is the same object in time and space. Despite the fact that a human being gets old, we cannot say that he/she is a different person in different times of his life.

For **The Infinity Problem** paradox, objects consist of an infinite number of properties which may have infinitely many values and hence are indeterminate. Therefore, the Identity of Indiscernibles is not usable because it is not possible to determine the identicalness of all the properties of two objects in order to determine their identicalness.

The **Ship of Theseus** Paradox is, in a way, related to the Paradox of Time and Change. It puts down the Indiscernibility of Identicals because the fact a human being had a heart valve replacement does not make him a different person.

There are other paradoxes worth to take a look as The Symmetric Universe, The Impact of Quantum Mechanics and The Paradox of Constitution. However, they will not be discussed here because it would be deepen to much the Identity matter for this thesis. For deeper insights on the Identity matter please refer to [12-13].

From the study of the Identity in Logic we can hypothesize that the better way of distinguishing identities or to check if they are equal, is not by verifying if they share the same properties or by admitting that they are equal, but rather to assure that they share the same properties except for their identifiers. Identifiers are what differentiate identities, so for two identities to be unlinkable in any context they must not share common identifiers.

2.2.3 Personal Identity

The identity of things is a controversial subject, as it could be seen on the previous subsection. However, personal identity is even more controversial and complex. Personal

Identity is about a person ascertaining the sameness of another person. It is about to be capable to discern one person from the other. There are several questions about the concept of personal identity and of course a variety of opinions asserting about it [12].

One of the opinions asserting about the personal identity of a person is related to the persistence of the person characteristics over time. It focuses on finding the requirements necessary to conclude that one person at one time is identical to another one at a different time. Due to the constantly changes of a human being though its life time it is not possible to determine how many differences are tolerable for stating whether it is still being the same person and not a different one in different periods of time. Therefore the qualitative identity cannot be used to determine the sameness of a person.

Another opinion states that the personal Identity is also strongly related to the psychological part of the human being. This psychological approach asserts that a certain psychological continuity is necessary for a person to persist. The human being inherits the mental features like beliefs, memories, preferences and capability for rational so, the present being mental futures are inherited from the past being. This approach meets a serious paradox. If it was possible to transplant the person's brain to one empty head it could not be longer said that the person remained the same after this change. A somatic approach to the identity term defines that identity is strongly comprised of some brute physical relation, the past or future of the being is related to the presence of the same body or biological organism through time. Both previously opinions assert that there is something that it takes for a person to persist. A more simplistic view of Identity states that a person existing at one time is identical with a human being existing at another if and only if they are identical. An opinion based on the evidences of the being states that the person here now is the one for some time before given the persistence requirements [12].

The different opinions asserting about the Identity term indicates that it is not yet fully understood and that human beings are still struggling to define it. No matter which definitions the human kind will reach for defining identity in the days to come, if an individual has a plurality of identities there must be a way that through one identity it is possible to reach the individual to which the identity belongs to.

2.2.4 Citizen Identity

To define citizen identity we have to do it within a nationality or government administration because citizen identity is citizenship. Citizenship is an attribute asserted by government according to its ruling. This part of identity is strongly physically oriented rather than psychologically, focusing more on the physical continuity of the person. Therefore, it is more adequate to use only a set of properties that are intrinsic to the physical body of the person and others extrinsic to it as family name, first name and address. These characteristics are called attributes and some of it can be also identifiers. Both can be temporary or persistent.

2.2.5 Real Identity

The real identity is used by us every day in the real world. It is the means by which people interact with the world and transmit their demands, wishes and transmit their intentions and emotions. The real identity is strongly attached to the physical presence of the identity, its extrinsic and intrinsic characteristics. It is also connected to citizen identity, as well as to the personal identity. The real identity is inserted within the context of the physical environment of the subject and is strongly affected by the physical presence of the body [7].

2.2.6 Digital Identity

Digital Identity can be defined as the digital representation of the overall known information about a subject across network systems. The subject can be a person, a group, a corporation, an organization, software, a machine, or any other identifiable entity. The digital representation of the identity is a digital collection of data representing the attributes, preferences, traits and claims of an identity. The information can be of any kind since it is related to and about the individual. The overall information can be a username and passwords, name, address, contacts, like email, phone numbers, mobile numbers, work address, IP address, bank account, written opinions, etc. Also make part of the digital identity the online identity defined next [10].

2.2.7 Online Identity

Online identity is part of the digital identity. It is a social identity that users establish in online communities. Online identity is outlined by the overall accounts users have in the various systems across the internet, such as the facebook social network, twitter, and youtube [10].

2.3 Representation of Digital Identity

A digitally mediated world requires that we express our identity through mediating layers of software design in order to reach our audience. There are thus two different parts of communications at stake: how humans represent themselves and how humans read the representation of others online. Within the constraints of computer-mediated communications, humans represent themselves through text or visual descriptions due to the lack of the physical presence of the body. Graphic avatars are often used within online games and forums to act out the intentions of their creators and express their identity online, even though most of the times they say nothing about the real body of the user. Music personality (sharing of music playlists) was also corroborated to constitute a tool of identity expression [15] and from it we may project a part of the identity of the person (its likes and dislikes) and we may even be capable to project a physical outlook in some cases. There are other kinds of identifiers that may tell a lot about the identity of a user, such as email addresses or mobile phone numbers. These are the most wide spread substitutes for the body. They are the virtual destination of the messages sent to a person and the source of the personal identity expression. Through them we may construct narratives about the owners. The domain of the email address may tell about the owner's features. For example, one of my emails is frank@ua.pt, the 'ua' domain, in this particular case, says that I am a current or former worker or student at the University of Aveiro, in Portugal. It may be checked just by reaching the domain www.ua.pt. The location of the university, the city as well as the country where it is located, dictates a higher probability of me being a Portuguese citizen than a Venezuelan one, even though both may be true. In the case of a phone country code, it may indicate a cultural context even if the owner does not live in such country. Digital representation may be even expressed to others unconsciously. A default name for the home wireless network may say the owner does not care about the security of his/her wireless network or he/she does not have the enough knowledge to configure it. Even though some of the identifiers mentioned here may represent the body, they cannot replace it. One attempt to ensure trust and consistency of the presence of the body is to introduce biometrics – to literally translate bodily identity into digital terms. In most cases, it is not a literal representation of the body that is being expressed, but a different representative form that substitutes it. Therefore, throughout the examples described, it can be said that the physical presence of the body cannot be replaced, just represented [7].

2.4 Digital and Real Identities Compared

From both digital and online identity we can realize that online identity is more about the social component of the digital Identity. It is about all the socials relationships across the Internet. In the case of the digital identity term, it is a more general concept. The digital identity encloses all the forms of information about an entity across online services providers which include the online identity, and tells about the owner's identity.

The real identity of a subject regards not only to the extrinsic attributes of the subject but also to the intrinsic ones. In the real identity, the physical part of the subject is strongly present and it is the means by which everyone contacts with physical world. The physical presence of the body is one of the most influential instruments of the real identity.

2.5 Next Generation of Digital Identity

In order to be possible to talk about the next generation of digital identity, a perception of how people see their identity in the real world must be taken because identity is something personal that may have ramifications on people's private life. What is acceptable to do and what is not must be known for projecting the next generation of digital identity in order not to go against the peoples' privacy requirements. According to the citizens' cultural context and its generations, some demands are more acceptable than others in what relates to people's identity intrusion. The people of North America have troubles accepting any identity document while they think that it is perfectly acceptable to have video cameras in every street corner or to apply DNA test to immigrants. In France DNA tests are unacceptable and were refused by the Chamber of Senators, video cameras are never welcome, but most people think it is normal for the police to ask for an ID Card whenever they want. Working in the identity field demands a study of the people's culture and history to verify what is and what is not acceptable to do in which is related to their identity [2].

One of the things people love the most is their privacy. Therefore this aspect must be taken into consideration when projecting the next generation of identity. In order to achieve total privacy, the users must have full control over the disclosure and dissemination of his/her digital identity and its information. Moreover privacy demands anonymity. Therewithal to fulfill these requirements mechanisms must be implemented to enable users to control the provisioning of their digital identity information to others. Also, the identity management systems responsible for the management of the digital identities must have means to avoid the linking between the digital identities and their holders. Trustworthiness is another requirement of the digital identity next generation. The information of the digital identity or a part of it must be true in order to make both the digital identity and the digital identity management reliable and responsible. No true information in a digital identity means no real identity behind it. Thereupon, several means have to be adopted in order to certify that at least some of the information about the user of the digital identity is true. The means by which these demands will be achieved have to be carefully analyzed because, according to the cultural context, a lot of people tend to suspect about technology applied to their identity [2]. The devices and mechanism adopted have to be trustworthy and flawless. That is why e-governance initiatives have a lot to contribute to the next generation of digital identity and to the identity management world by providing digital identities to internet users as they do for their citizens. E-governance may either be the step to make identity management transparent and it may also help on freeing governments from corruption [16]. E-governance could be the step to improve digital identity management on the Web, as well [17]. As can be realized, the technology needed to fulfill these demands is complex (and complexity must be avoidable), mostly for users rather than for the systems, because users are the main key to its propagation and diffusion. Users are the ones that say whether it is acceptable or not the use of some kind of technology in their daily life. For the mentioned responses, the next generation of digital identity must considerably care about each particularity of people's identity when projecting any new digital identity solution.

2.6 Personal Digital Certificate

Digital certificates are part of a developing set of technologies that can address many of today’s security, identity and accessibility issues. They are part of a technology called Public Key Infrastructure (PKI). Digital certificates are made by a cryptographic key pair and information about a user, as well as a certain collection of attributes of the subject. Therefore, digital certificates have been described as virtual ID cards. Each certificate is digitally signed by the issuer authority. The cryptographic keys asserted to the certificate may be used to sign, encrypt and decrypt documents. Consequently, it may be used for authentication purposes, as well. The public key, as it says, is available to others. It checks the signature and encrypts documents. Otherwise, the private key is kept secretly to the holder possession, and it is used either to sign or decrypt documents. Digital certificates are asserted or created by Certificate Authorities (CAs). These authorities are responsible for issuing, managing and revoking digital certificates. An example of a CA includes VeriSign, a well-known commercial provider[18].

The following table provides an illustration of a certificate composition.

Table 1 - Certificate structure.

Certificate version n.º	
Certificate contents	Private information
	Private information encapsulated
	x. 509 certificate
	Certificate cancelation list
	Others
	Certificate sub contents
Certificate content and encryption key	
Certificate finger mark	

Certificates may be of many types. The most relevant are:

- Root or authority certificates – are certificates that create the base or the root of a certification authority hierarchy. These certificates are not signed by another CA – they are self-signed by the CA that created them. When a certificate is self-signed, it means that the name in the issuer field is the same as the name in the subject field.

- Browser Certificates - used for securing and authenticating the communications between web browsers and web servers. This type of certificate is the most popular.
- Client Certificates - this type of certificates are also known as end-entity certificates, identity certificates, or personal certificates. They are certificates that an entity holds for proving its identity when accessing a certificate enabled web service. The use of certificates in theses situation are correlated to mutual authentication of the service's session participants. In order to prove the identity by using the certificate, it must belong to the chain of certification in the first place.

The parties involved in a certificate cycle are:

- The issuing party – the party that digitally signs the certificate after creating the information in the certificate or checking its accuracy.
- The requesting party – the party who needs the certificate, and will use the information of it for identification or access purposes.
- The verifying party – validates the signature on the certificate and then relies on its contents for some purpose. It is the one that is going to consume the certificate.

Digital certificates together with smart cards have proven a relevant value for identity confirmation and business negotiations legal binding [19].

2.7 Smart Cards

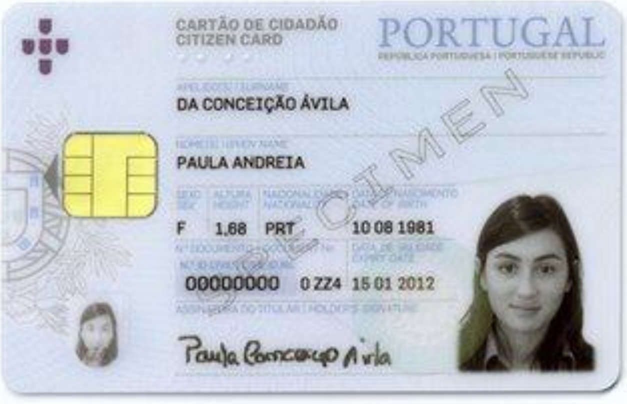
Smart cards are portable tamper-resistant cryptographic devices that play a key role in digital identity by securely strong the card owner identity attributes and preserving its privacy, and by providing strong authentication of the card owner before releasing identity attributes [20]. A huge number of smart cards are deployed by mobile network operators to authenticate and identify subscribers to the GSM and 3G networks, and by banks and financial institutions for payment. Large deployments are also on the way for government identification cards or electronic passports. The security of the smart cards against physical and logical attacks has been achieved thru the development of advanced counter-measures and as a result, smart cards are the *de-facto* standard for digital security. Smart cards are able to hold certificates, becoming a mean for strong authentication, which can provide a two-factor authentication, i.e. something I have (the smart-card) and something I know (the pin of the smart card). The smart cards are being implemented with PKI features as means of citizen identification, authentication and signature, and for the access of a wide range of online state services. Several European countries have implemented smartcards and certificates for their official citizenship identification documents. There is even a European project to establish a European eID interoperability that allows citizens to establish new e-relations across borders, just by presenting their national eID (see www.eid-stork.eu).

2.8 Identity proof

The proof of identity is done by claiming to have some attributes. Usually, identity proof is done by presenting a document provided by an entity, in the case of citizenship, the citizen cards are the elements asserted for that purpose.

2.9 The Portuguese Citizen Card

The Portuguese Citizen Card is a convergence of technologies and government capabilities. It is inserted within a European e-government initiative. The Portuguese Citizen Card is the official document given to the Portuguese Citizens to enable them to securely identify themselves to others by validating their claims with a reliable document given by a civil registration institution under the administration of the government of the Portuguese Republic. As a technological document it allows its' holders to identify themselves when dealing with computerized services and to authenticate electronic documents.



CARTÃO DE CIDADÃO
CITIZEN CARD

PORTUGAL

DA CONCEIÇÃO ÁVILA

PAULA ANDREIA

F 1,68 PRT 10 08 1981

00000000 0 ZZ4 15 01 2012

Paula Conceição Ávila

Figure 1 - Illustration of the Portuguese Citizen Card.

The Citizen Card has printed in both sides intrinsic information about its holder. The information that can be seen in both sides of the card is:

- A picture of the users face.
- Name, family name and the name of the parents.
- Nationality.
- Civil ID number, taxes number, social security number and health number.
- Gender, height, date of birth.
- Date of expire, holder' signature.

The Portuguese Citizen Card has an embedded smart card to enable Portuguese Citizens to identify themselves when interacting with the State's electronic systems. This smart card holds the information printed in the card, a digital certificate for electronic authentication and signature purposes, digital finger print information and other more information as the holder's current address [4].

2.10 Laws of the Digital Identity

The Laws of Identity referred here are the ones made by Kim Cameron, Architect identity at Microsoft and author of the blog www.identityblog.com. These laws explain the successes and failures of digital identity systems [21]:

1. User control and Consent:
    a. Any personal information should not be revealed without the user's consent.
    b. The user is in control of what information is released and what digital identities are used.
    c. The systems should warn the user about the consumers systems of his/her identity and that it goes to the right place.
    d. The user should be aware of the purposes of collecting their personal information.
2. Minimal disclosure for a constrained use:
    a. Systems should implement principles of limited information to avoid risks.
    b. Minimize aggregation of identity information to minimize the risk.
    c. Systems shouldn't use unique identifiers.
3. Justifiable Parties:

   a. The user should be aware of the parties or parties with whom he/she is interacting while sharing information.
   b. The disclosure of identity information is limited to parties having a necessary and justifiable place in giving identity relationship.
   c. Every party to disclosure must provide the disclosing party with a policy statement about information use.

4. Directed Identity:
   a. A universal identity system must support both “omni-directional” identifiers for use by public entities and “unidirectional” identifiers for use by private entities, thus facilitating discovery while preventing unnecessary release of correlation handles.
5. Pluralist of Operators and Technologies:
   a. A universal identity system must channel and enable the inter-working of multiple identity technologies run by multiples identity providers.
   b. A universal system must embrace differentiation of contexts (roles).
6. Human integration:
   a. The universal identity systems must define the human user to be a component of the distributed system integrated through unambiguous human-machine communication mechanisms offering protection against identity attacks.
7. Consistent Experience across contexts:
   a. The unifying identity metasystem must guarantee its users a simple, consistent experience while enabling separation of contexts through multiple operators and technologies.

Even though these laws were made specifically for the Microsoft metasystem, they should be taken in account when projecting any digital identity management system.

3 Digital Identity Management

The definition of Digital Identity Management differs according to the context in which it is seen. For the purpose of this master thesis Digital Identity Management is defined as a discipline that consists of processes, policies and technologies to manage the complete lifecycle of the user identities, the disclosure and dissemination of its information across the services online as well as to control the user access to the systems resources (services providers) by associating user rights and restrictions. These resources include information services process capability, buildings and physical assets. Identity management can be also referred as Identity and Access Management [22].

Throughout this section it will be mentioned, described, and explained the digital identity management discipline. Firstly, it will be introduced a variety of definitions involved in the digital identity management field, and then it will be approached the problems of digital identity. The authentication subsection will described the main authentication methods. Further a key element of digital identity management, the Public Key Infrastructure and the application protocols for the transportation of identity information: the HTTP and TLS will be introduced and described. Last, it will be made the section analyses in the discussion subsection.

3.1 Concepts

In order to understand better the Digital Identity Management matter and its models a few concepts must be firstly introduced [23-25]. The definitions of the section 2.1 must be considered

User – the term ‘user’ refers to the physical person who is interaction with the computer system. The user is a virtual entity for a computer system that is represented by a persona (or a collection of personae).

Persona – the term ‘persona’ is used for the digital representation of user’s characteristics. The user may maintain several personae that may be more or less related to each other. The characteristics of a persona are represented in the form of attributes. Thereby the personae are part of the digital identity of the user.

Account – an Account is a data structure that is usually kept in the computer system databases. The account is used for access control purposes, storing attributes, credentials, etc. Account is usually used as a persistent storage for (partial) persona attributes, but it also may be unrelated to any physical user persona.

Entitlements – are the name for the services and resources to which an identity is entitled: credit limits, disk space, bandwidth allocations.

Permissions – are the actions that the subject is allowed to perform with respect to the resource: withdrawing funds, completing a purchase, updating a record, and so on. Permissions are related to access control ruling.

Credentials – are the elements to prove that a subject has the right to assert a particular identity and a way of transferring trust between entities. Credentials may be a username and password, a X.509 certificate or biometric data.

Security – is the protection of the user's information integrity and its unauthorized access. Security can be seen as an authorization issue.

Biometrics – refers to the recognition of human by the analyses of their intrinsic physical and/or psychological traits with the purpose of determine or confirm the identity of an individual.

Identity – is specified by the set of permanent or long-lived temporal attributes associated with an entity. Attributes are what characterizes the identity.

Identity Provider (IP or IdP) – is an entity that provides digital identities to users and to whom the users entrusts the ability to make assertions about it. The IP is responsible for the management of the identities information according to his/her self needs. The service providers that require related information about a user must contact the IP of the user in order to obtain that information. Since the identity provider is the entity that makes assertions regarding the user's identity a minimal amount of trust between the holder of the identity and the identity provider is assumed to exist.

Self-Identity Provider – refers to the sources of information that resides in the user system (laptop, P.C.). One example of self-identity provider is the Windows Card Space by Microsoft.

Service Provider (SP) – is an entity that provides services to other identities. Usually providing subscription based service to individuals. Examples of service providers are: Google Services (Email, Docs, Picasa, Gtalk, etc.), Facebook, MySpace, LinkedIn, Youtube, Flirk.

Identity Silos – was the name given to the fortresses of digital identities of the most used services of the Internet. This name has origin on their restriction policy to the access of the information. The YahooID is an example of an Identity Silo.

Identification – is established via authentication. In terms of user identities identification is the process through which one presents him/herself as the owner of the identifier by presenting credentials associated with that identifier. For example when one provides his/her identity card at a request of a police man.

Registration – users adhere to services by registering in them. The registration establishes a bound between the user that is registering at the service and the service provider itself. In these situations users provide the services with pieces of information about them, such as email address, date of birth and depending on the purpose of the service, home address or any other shipment address and mobile may be requested, as well.

Authentication – is the process of positively verifying the identity of an entity, as a prerequisite to allow access to specific resources in any system. During authentication process an association between an identifier and individual is certified by the individual's credentials. For example, an automobile is identified by its license plate, and that is authenticated as legitimate by the database of cars that are not being sought for enforcement purposes. Identification and Authentication are related subjects.

Authorization – is related to access control. It is a decision to allow a particular action based on an identifier or attribute. Examples include the ability of a person to make claims on lines of credit, the right of an emergency vehicle to pass through a red light or a certification of a radiation-hardened device to be attached to a satellite.

Identity Identification – is obtained by proving the association between an individual and an identity. For example: the association of a person with a credit or educational record.

Attribute Authentication – is enabled by proving the association between an entity and an attribute. For example: the conformation of the age of an individual. This is usually a two-step process, where the association between an entity and an identifier is established, and then a link between identifier and attribute is established.

Single Sign On (SSOn) – is used to refer to an authentication method, that enables users to only present their authentication credentials once while gaining legitimate access to several different services. Up until recently this was common among service providers within the same domain, for example inside a company an employee would have to authenticate only once and would immediately be authenticated to all the necessary company services (Centralized identity [Active directory, Kerberos]). The SSOn role is to securely transfer the user’s identity, attributes and current authentication status from source site (Identity Provider) to the destination site (Service Provider) seamlessly. Since the user relies his/her authentication and assertions regarding his/her own identity to the identity provider at least a minimal relation of trust is assumed to exist between the user and it.

Anonymity – is defined as the inability to identify an entity within a set of attributes. In order to enable it for an individual, a set of others individuals must exist with the same attributes, essentially making an individual indistinguishable from a set of others. The larger the anonymity set is, the harder will be for an attacker to distinguish an individual from the remaining individuals in the set. Ideally the anonymity set will be such, that the attacker will not be able to sufficiently single out an individual. Anonymity can be obtained with the use of pseudonyms (Pseudonymity).

Pseudonymity – refers to the use of pseudonyms as identifiers. A pseudonym is an identifier that has no direct relation to the user’s real identity; it is used only to refer the identity of the user in one or more contexts. Pseudonyms are usually employed for avoiding the use of the user’s name or other attributes related to the user. They are used so that the real world identity behind the pseudonym is only know to the identity, and otherwise is hidden to all other parties. Pseudonyms belong to a specific persona and are typically long-term identifiers. The persona to a pseudonym mapping is in most cases private information, the persona to which a specific pseudonym belongs to is not publicly known. Pseudonymity is somehow related to unlinkability (Indirect Linking).

Linkability – is defined as the ability to determine whether two elements are related or not. In terms of user identities, linkability is the ability to determine whether two identities are related to the same individual. Within the context identity management, two identities are linkable if they share some sort of attribute that allows one to conclude that both identities belong to the same individual. For two identities to be not linkable they must not share any kind of (unique) identifier or attribute that would allow that sort of conclusion to be inferred.

Privacy – is related to the protection of the attributes, preferences, and traits associated with and identity from being disseminated beyond the subject's needs in any particular transaction. It is built upon a foundation of good information security and that is dependent upon good identity management. Privacy is related to anonymity and usually it is controlled by a privacy policy.

Identity Privacy – is the ability of an individual to seclude him/herself about his/her own personal information. In order to an individual have privacy he/she must have full control over the disclosure and dissemination of his personal information.

Self-Identity Provider – is an identity provider owned by the user, it may be stored in one of his/her own systems, such as in a laptop. The Windows Card Space application (discussed later) is one example of a self-identity provider.

Privacy policy – is a common set of rules to express privacy practices. It is employed by one for one or more individuals. Privacy policies are the first step towards informed consent for data collection. Such ruling should include:

- Which user information will be collected.
- The purpose given to the collected information.
- The recipients of the collected information.
- User's rights to access and modify the collected information.

Usually the users tend to see the presentation of privacy policies by services providers (SP) as a sign of good faith. However, there is no practical solution to verify if the services providers actually adhere and comply it. Usually the most service providers state that no information about the user will be collected without the user consent. Nevertheless, they are not clear about what will be the intended use for the information gathered, especially when it comes to its disclosure to third parties. Even when they claim the user information will not be disclosed without any consent, they contradict themselves by stating that the information will be always shared with their business partners. This kind of loopholes in the privacy policies indicates that the user cannot certainly control the collection, disclosure and proliferation of his/her information with accuracy, even if his/her service provider complies with his/her privacy policy. Moreover web applications work using a client-server mode, where the user initiates the interactions. This makes some of his/her network identifiers (L3 and perhaps L7) available to the service provider with or without the user consent.

3.2 Problems of Digital Identity

3.2.1 Identity Theft

One of the mains digital identity drawbacks is the digital identity theft. Digital Identity theft is on the rise, affecting almost ten million victims in 2008 (a 22% increase from 2007) [26]. Online there are around 55.000 phishing websites worldwide where only in the U.S.A are located around 43% of them. Only in the United States in 2007 $3.2 billion was lost due to identity theft via phishing [27]. In the countries of the European Union the proportion of users who have been victims of phishing remains below 10% in each Member State, ranging from 8% in Malta and 7% in Ireland to only 1% in France, Sweden and Slovakia. Most of these cases happen due the lack of awareness as well as the general reluctance to admit that one has been fooled. There are not available statics on total fraud in the EU but estimates in relation to card fraud in the EU were between $500 and $1000 Million, and it is not necessarily decreasing. More worrying, payment fraud is increasingly taking a trans-national nature [28]. There are many websites selling personal data as credit card numbers, expiry dates, the CCV code and other security information [28]. These mentioned problems strangle the privacy as well as the anonymity required for a digital identity.

3.2.2 Lack of Trustworthiness

Although the lacks of the digital identity mentioned in the last sub subsection are important there is one more imperative: lack of trustworthiness. It exists because the information that shapes the digital identity is not acquired and confirmed by accurate and reliable means, therefore, it can be false and normally most of the internet systems do not try to validate it. Along with the lack of digital identity certification comes the lack of accountability, e.g.: on forum or news comments anyone can write whatever comes to his/her mind without taking responsibility for it. The lack of digital identity responsibility has caused a lot of troubles what have been causing huge costs to many internet systems. Under this lack of reliability and responsibility are billions of digital identities across internet systems. A solution to overcome the lack of accountability of the user and therefore credibility the digital identity is by the acquisition of some of the user's physical features. The cost-benefit and the non-user-friendly issue of the devices needed did not permit the wide diffusion of this mean Online [29], even though they are effective, sufficient and the requirements to reinforce trustworthiness online. Some of them go against the users' identity requirements, such as their privacy reclusion.

The lack of trustworthiness of the identity information and the lack of control over the disclosure and dissemination of identity data are the main drawbacks of the digital identity management systems since ever. The digital identity management field of development is of the most importance nowadays by IT companies due to its impact on the eBusiness, the eGovernment, the Homebanking and the overall Internet Economy as can be realized by the following quotation:

> *"We declare that, to contribute to the development of the Internet Economy, we will... strengthen confidence and security, through policies that... ensure the protection of digital identities" [30].*

3.2.3 A Culture of Privacy

Privacy is important. Privacy concerns every user online, or at least it should. The loss of privacy may lead to the loss of anonymity, and it may carry undesirable issues, such as Identity theft. Most often users think that the privacy matters online are of the responsibility of systems and IT experts. However that is not always the case. A culture of security must be implemented by the users [31]. They are the owners of the information. Even though they are not responsible for the security of the channels of communications, they are responsible for the provisioning of their personal information to the systems online. Therefore, they are the first agents to check whether providing the information to a specific system is acceptable and free of risks. They should always check whether the site they are visiting is of trust or not. Many identity attacks succeed because the user was fooled by something presented on the screen, and not because of insecure communication technologies, such as is the case of phishing attacks. Most of them occur not in the secured channel between web servers and browsers — a channel that might extend thousands of miles — but in the twenty or thirty centimeters between the browser and the human who uses it. Hence, users should take preventive measures regarding their private information in order to avoid actions that may lead to undesirable and uncomfortable situations [32]. One of these preventive measures users may take is to check if the system is trustable by checking the identity information provided by the browser he/she uses to access the Web as illustrated further.

3.3 Authentication

As described before, authentication is the act of proving something. Usually this process is made by presenting claims. Next it will be indicated and briefly described the authentication factors, and the most used authentication methods on the Web.

3.3.1 Factors

Authentication methods depend on human factors divided in three classes:

- Something one knows (a password).
- Something one has (a hardware token).
- Something one is (a finger print).

Recently, a fourth-factor of authentication was proposed to also be taken in account:

- Somebody you know [11, 33].

This last factor is based on vouchering, trustworthy and tight social relationships. When a user, for instance, lost is password, another user of the system, sends him a voucher for enabling him to temporally authenticate in the system while he/she has not his/her authentication process ratified. This authentication method is more suitable to use within institutional contexts rather than in a public service online since the trustworthy bound is probably stronger.

3.3.2 Methods

Password based Systems

Password based authentication is a simplistic method of authentication, and is the most widely use form of authentication online. It has been a cornerstone of computer security for decades. Password based authentication is an efficient mean for sharing a secret between a user and a system. In order to implement a password based authentication method it is not required any special software on the users' computers. Passwords are portable, users only have to memorize them, and it authenticates the user directly because only the user knows the password [34].

Password based authentication has many drawbacks. Strong passwords are not easy to remember and users tend to write them down, making them vulnerable to password thieves. When users share passwords the ability to audit and trace access to a particular user is less possible. Passwords may be known by the administrator of the system, and he/she may use it to discern the user. Plus, a single password may not be safely used in each affiliated web application, due to password format restrictions. Passwords are easily guessed and when systems accept it as plain password (seed) that is chosen by the user and do not implement neither delays response (account locking) nor captchas, they are easily vulnerable to dictionary attacks [35-36]. Furthermore, passwords sent in clear text over the network, are easily subject of eaves dropping and replays. They may even be manipulated on the exchanging channel, when no cryptographic transport protocol is implemented (man-in-the-middle attacks). There are several proposed schemes to overcome the issues of the simple password authentication, such as one time password and challenge-response. However, when used in a plain TCP/IP (no TLS or SSL) environment connection, the data can be manipulated after a successful authentication takes place. Nevertheless, it may be avoided by pre-authenticating the server to the client (mutual authentication), by using other independent methods or explicitly authenticating the transported data by digital signature confirmation [25].

Hardware Tokens

Password based mechanisms may be strengthen by the use of hardware tokens. Hardware tokens are an authentication example of "something the user has" [37].

Benefits:

- The use of tokens prevents a thief with a stolen password from accessing the web site, because he/she would have to steal the physical token from the victim as well.
- They prevent the sharing of accounts since they would have to be duplicated.
- Hardware tokens are portable.
- No special software is required on the user's computer.

Drawbacks:

- Tokens are expensive and must be replaced or refurbished every few years.
- Token are easy to misplace or damage. A lost token prevents a valid user from accessing the web site, which disrupts business or commerce.
- Tokens are inconvenient since the user must manually enter the value of the token as well as the password.
- Some token devices may not be of affordable price for the end user.

Biometrics

Biometrics fits on the "something one is" authentication factor. It represents physical (intrinsic) characteristics of a human being and depends on its biological individuality. Biometrics have the following characteristics [38]:

- Universality – Generally, every person should have the same type of characteristic. People, who have not, as mutes and handicaps, must be accommodated in some other way.
- Uniqueness – Generally, no two people have identical characteristics. However, identical twins are hard to distinguish.
- Permanence – The characteristics should not vary with time. A person's face, for example, may change with age.
- Collectability – The characteristics must be easily collectible and measurable.
- Performance – The method must deliver accurate results under varied environmental circumstances.
- Acceptability – The general public must accept the sample collection routines. Nonintrusive methods are more acceptable.
- Circumvention – The technology should be difficult to deceive.

Typically used biometrics characteristics for authentication are: finger print, palm print, voice, face and iris. They are automatically extracted at the time of the authentication process and compared with samples of the same elements kept in a database. Even though the use of biometrics is a hope in the authentication mechanisms of the future [39], it shouldn't be used alone because they are associated to human characteristics, and so they tend to change through his time life or may be even temporally affected, such are the cases of a person's face and a person with hoarseness, when facial and voice recognizers devices are in use. A combination of elements or factors should be used together according to their possible unavailability [11, 40]. Biometrics can be used for either positive or negative identification, e.g. a worker is allowed to access a particular area in a company (positive identification) or a particular person is not in the watchlist of wanted people (negative identification) [41].

Benefits:

- Authenticate a user through a unique physical characteristic.
- Directly authenticates the person, not indirectly through a password or token.
- Biological features are difficult to steal, thereby they make biometric authentication very strong.
- Biometric feature is eminently portable and is unlikely to be lost.

Drawbacks:

- It is necessary appropriated biometric hardware and software.
- Some biometric readers may not be of affordable price for the end user.
- False positive (allowing an invalid user).
- False negatives (forbidding a valid user).

3.4 Public Key Infrastructure X.509

The Public Key Infrastructure X.509 (PKI) is the most widely used public key certificate format in both enterprise and Internet environments. The PKI is made by a set of methods and formats for the management of encrypted public keys and related data. The foundation of the Public Key infrastructure (PKI) is a pair of mathematically related asymmetric keys. One key can be used to encrypt a message that can only be decrypted using the other key. It is computationally infeasible to discover one key by knowing the other. The public key as it says is published to the world while the private key is kept in a secure place. The public key is used for encryption and signature verification of data while the private key serves for decryption and signing data. Each entity's identity in the PKI is bound to a digital certificate by the Public Key Certificate (PKC) issuer. A PKC is an electronic data structure that contains the entity's identification, issuer identification, a certificate describing data, and any other relevant data bound to the public keys values of the identity. See subsection 2.6 for a more described explanation of a PKC. PKCs are used in many communication systems on the Internet for authenticating the communicating agents (mutual authentication) and to encrypt the data in the exchanging channels [25, 42]. The most common communication protocol in use today that employs X.509 public key certificates is the Transport Layer Security (TLS) Protocol [43].

A simple public key infrastructure starts with a certificate Authority, which issues the certificates to the end user. The PKC of an entity is always signed by the PKC issuer's (CA) private key. More complex PKI can be established by the intermediation of more than one CA. In this case one CA must be the root CA, which owns a self-signed certificate and issues certificates to subordinated CA, who in turn issue certificates to the end users. In this case the certificates are issued in a form of tree structure, where the Root CA's certificate is on the top of the tree (root) and the subordinated CAs are further down the tree and inherit the trustworthiness of the root CA certificate. Each certificate on the chain depends on the trustworthiness of the previously certificate in the chain. Issued certificates have a specific period of validity. Therefore, from time to time, the CAs publish a Certificate Revocation List for revoking the issued certificates before the end of the validity period. The Online Certificate Status Protocol (OCSP) is protocol that may be implemented by the CA to provide the revocation status of the issued certificates online.

The several parts that make the PKI have drawbacks, and it should be followed a guide line of recommendations on its implementations [43]. Certificates should only contain minimal information about the subject, since once the certificate is present, all the attributes and data in the certificate are disclosed to the verifying party. Additionally the subject's privacy may be violated, since his/her/it activities at different sites may be correlated by his/her/its public key.

The PKI X.509 became an important element on Digital Identity Management.

3.5 Transport Protocols

3.5.1 Transport Layer Security Protocol

The Transport Layer Security (TLS) is the successor of the Secure Socket Layer (SSL) by IETF, a protocol that provides security for communications over networks, such as the Internet, by the implementation of cryptographic techniques. It is not necessary appropriated software in order to establish a TLS session between two end points. The TLS protocol is an application layer protocol (OSI Model), it runs on top of the TCP/IP protocol (transport and network layers) and encrypts its segments. The data is only protected while it is being exchanged in the channel; if data is stored on the end points they are no longer protected. The TLS does not depend on any specific cryptographic algorithm but rather on a set of them pertaining to key exchange, encryption, hashing, and digital signatures. A Cipher suite of cryptographic algorithms is agreed after a TLS session has been established by the end points of the communication in order to protect the data exchanged between them.

The TLS is made up of two layers:

- The **TLS record protocol**: encrypts the data to be exchanged.
- The **TLS Handshake protocol**: deals with the authentication of the communicating agents and with cipher suite agreement.

The protocol supports several authentication modes:

- **Total anonymity**: In the total anonymity mode there is no authentication of the communicating parties. This mode is included for backward compatibility only.
- **Authentication server**: Only the server authenticates itself to the client. The client by itself remains anonymous.
- **Mutual authentication**: Both parties in the communication authenticates to each other by exchanging their respective public key certificates and appropriate proofs of possession of the private keys.

The TLS protocol has several drawbacks. It allows the discernment of the parties involved in the communication because the certificates are exchanged in the clear [44]. The use of TLS generates an increase of the connection overhead and it may decrease the speed of the connection. The TLS protocol is wide used to secure the WWW, by serving as a secured channel of communications for the HTTP protocol (HTTPS). The IETF is redefining the TLS as a high priority (http://tools.ietf.org/wg/tls) due to a recently discovered problem on the TLS handshake mechanisms that may lead to an attack (man-in-the-middle) [45].

3.5.2 HTTP

HTTP is widely used on the World Wide Web for transferring hyper media information between systems that usually runs on top of the TCP/IP protocols, usually on port 80. Usually the agents of an HTTP sessions are a client (user's browser) and a server (the data source).

The HTTP uses Uniform Resource Locators (URLs) a subset of Uniform Resource Identifiers (URIs) for the identification of the resources to access. The format of an URI (defined by http://www.ietf.org/rfc/rfc2396.txt) is the following:

- <scheme name> : <hierarchical part> [? <query>] [#<fragment>]

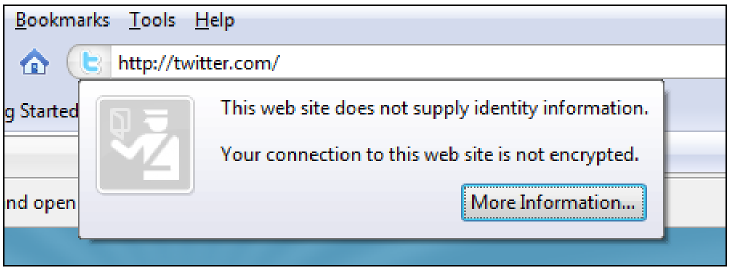
The hierarchical part is separated in <authority> and <path> as showed next.

- <scheme>://<authority><path>?<query>#fragment.

The <scheme> tag represents the protocol being used. The <authority> tag represents the server. The <path> represents the path to the resource <authority> (server). The question mark represents the introduction of a query and the # represents a particular secondary resource (such as an html macro). Resources communicate by making requests. The main HTTP request formats are:

- GET – a GET request, as the name says, is performed for getting HTML formatted data from a system.
- POST – the POST request is used to submit data to be processed by the system to which the post is effectuated to. Usually these results are new resources, update of resources or deletion of existing resources.

The HTTP exchange data on the clear, therefore in order to protected the data transmitted, the HTTP messages must be send over a secure channel HTTPS (HTTP/TLS). When a communication between agents is being established over an HTTP secure channel, the URI scheme usually shown is HTTPS. Users may know whether the authority they are reaching is secure by looking to the certificate warning of their browser in the address bar. Figure 2 shows it on the Firefox Web Browser.

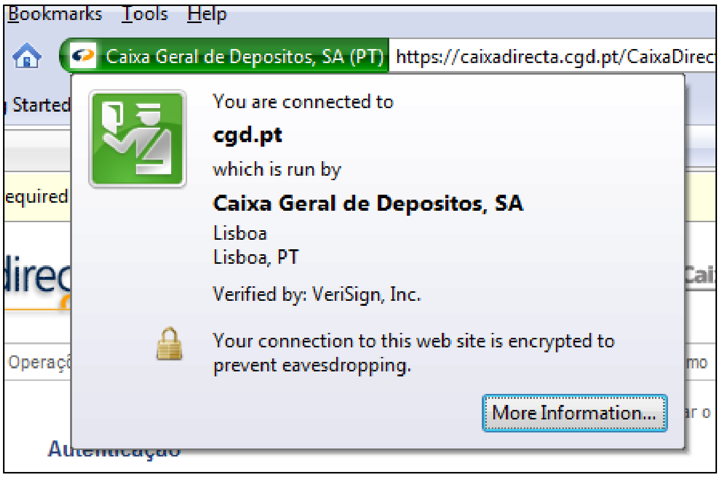


This web site does not supply identity information.

Your connection to this web site is not encrypted.

Figure 2- Certificate information of Twitter.com.

On the Figure 3 it is shown that the communication with the specific website is not secure.



You are connected to
cgd.pt
which is run by
Caixa Geral de Depositos, SA
Lisboa
Lisboa, PT

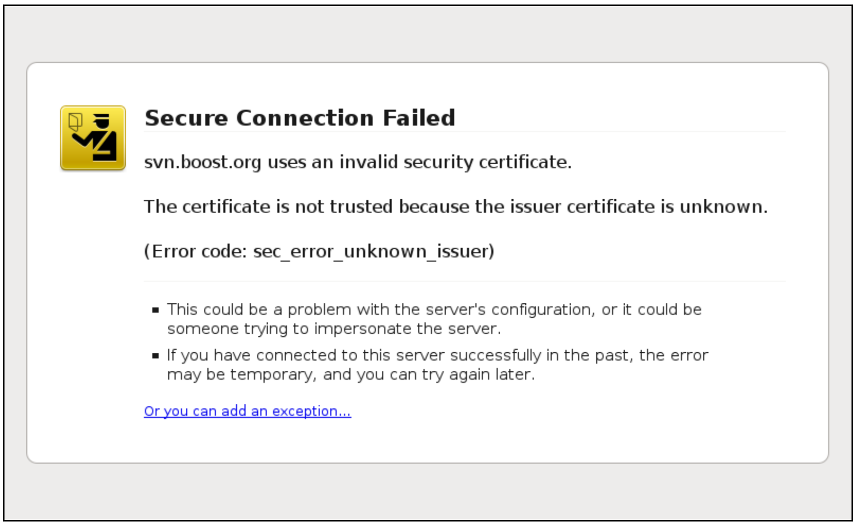
Verified by: VeriSign, Inc.

Your connection to this web site is encrypted to prevent eavesdropping.

Figure 3 - Certificate information of CGD.

Figure 3 shows the identity information of the digital certificate made available by the service. This way the user may check whether accessing the specific server is secure or not by looking at the available information provided by the service's certificate. As it is indicated by the information window, the connection to the web site is encrypted to prevent eavesdropping. Therefore attempts from an eavesdropper to read the data exchanged between the browser and the server will be unsuccessful. The Certificate Authority that provided the certificate is also shown. It is the well-known commercial CA VeriSign, Inc, which indicates that the entity that holds the certificate is reliable.

The scheme of the URI is HTTPS, which also indicates that the connection with the server is made over a secure channel, probably TLS. In such case the TCP port used will be 443 and not 80. When the website is not trustful, firefox display a message similar as the shown by Figure 4:



Secure Connection Failed

svn.boost.org uses an invalid security certificate.

The certificate is not trusted because the issuer certificate is unknown.

(Error code: sec_error_unknown_issuer)

- This could be a problem with the server's configuration, or it could be someone trying to impersonate the server.
- If you have connected to this server successfully in the past, the error may be temporary, and you can try again later.

Or you can add an exception...

Figure 4 - Unsecure connection warning from the Firefox Web Browser.

In order to understand how HTTP and TLS are displaced in the layer of communication, the following figure is shown.

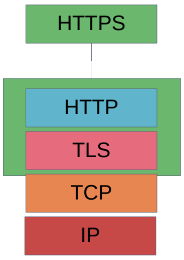


Figure 5 - The displacement of HTTP and TLS on the communication layer.

In the Figure 5 it is illustrated that the HTTPS protocol is defined by the HTTP protocol running on top of the encrypted transport layer security. Thereby it is possible to securely transfer the information carried in the systems communication channels because the TLS segments are encrypted by the protocol.

3.6 Discussion

In this section it was provided an approach to the digital identity management. Some definitions implied in the specifics of digital identity management were presented. It was also pointed out that users are the first responsible agents on the provision of their particular data since they are the data suppliers. Several authentication factors and methods were indicated and described as important elements of a digital identity management system. It was told that the authentication method to implement in an identity management system should be chosen according to the security context required by the system and that several systems failures occurred due to the weakness of the system authentication method. More efficient means of authentication were introduced, such as the use of smart cards with digital certificates and biometrics (hardware tokens) as a possible means of implementation to overcome the authentication issues.

The Public Key Infrastructure is with no doubt an important mechanism for the management of digital identities. It implements mechanism that provides authenticity to the identity and allows the actors to mutually identify themselves legitimately. The PKI is also a means to prevent identity theft. The implementation of a PKI also protects the integrity of the messages exchanged between systems by the use of a signing mechanism and provides security against eaves droppers in the channel of communication by implementing encryption techniques.

The key protocols for the transportation of identity data and mechanisms to inform how a user may know if the connection he/she is establishing with the service is secure and whether the service is a reliable one were also indicated.

Next it will be introduced the different models for managing digital identity information.

4 Digital Identity Management Models

Identity Management is traditionally seen from the service provider's point of view, meaning that it is an activity undertaken by the service provider to manage the users' identities. Service providers also took as granted that users need to authenticate to them, otherwise they do not need to authenticate themselves to the users. This drawback is the cause for the many phishing attacks Online that have been causing a lot of costs to the services providers, users, and business companies working Online [46-47].

The lack of an Identity Layer in the Network ISO framework with the dynamization and the user customization of web applications forced the creation of systems to manage users' identity [3]. Several models were created to tackle the management task.

Firstly in this section it will be introduced the relations definitions implied in the digital identity management specification. Throughout, it will be mentioned, described, explained and analyzed the various models used by the millions of services providers online for the management of their users' identity data. The objective is to know their benefits and pitfalls in order to project a digital identity management solution that can benefit both users and service providers, and avoid the drawbacks of the traditional identity management. This new digital identity management solution will be introduced, described, explained, analyzed and compared with other existing identity solutions in the subsection number 8.6 of this master's thesis.

In order to understand the identity models, it should be well understood first the relation between the different terms implied in the digital identity management subject. The most important terms are: entity, identity, characteristics or attributes, identifiers and unique identifiers. Further on this thesis others terms will be introduced. But for now let's observe how the already introduced ones are associated.

Figure 6 illustrates the relation between entities, identities, characteristics / attributes and identifiers. An entity can be any institution, device or person, usually called subject. Each entity is an identity and each one has several characteristics. These characteristics may be identifiers for identities, and some of these identifiers may be unique identifiers, as an Email and a Civil ID number are. Each one of these entities is associated to one or more identities that in its turn are associated to several own characteristics / identifiers.

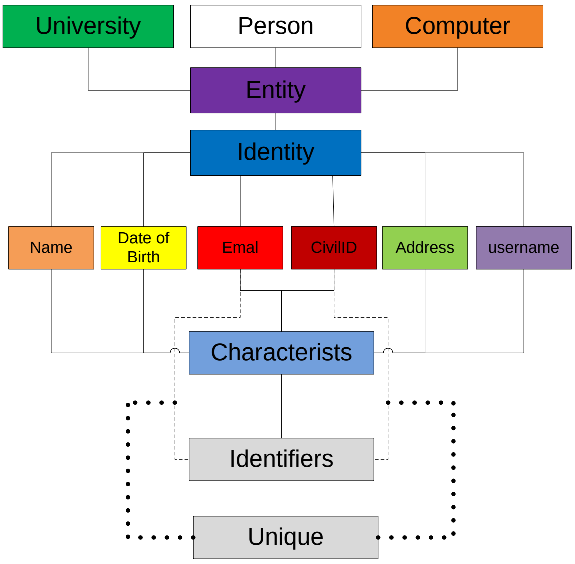


Figure 6 - Identity Scheme.

4.1 The system-centric model

The system-centric model was the first solution to appear for managing users' access to a service provider. It is called the traditional model of identity management because it is widely used by the online services providers for a long time. In this model the service providers act as both identity provider and service provider in an united working block [48]. This line of work isolates service providers from each other's because users get separate credentials to each service they want to use.

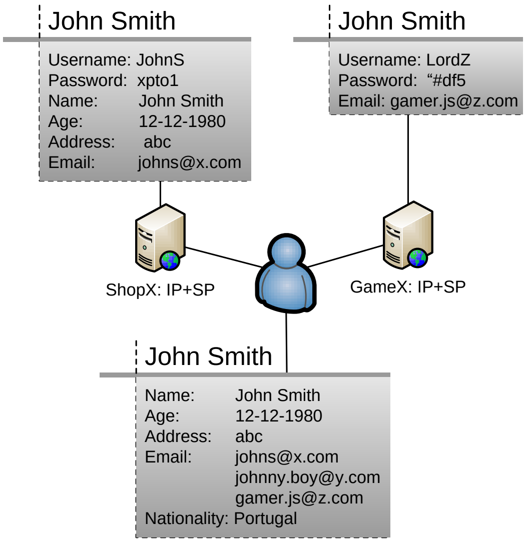


Figure 7 - Traditional Identity Management Model.

In Figure 7 it can be seen that every system compounds the IP (Identity Provider) and de SP (Service Provider) in a united working block. The user has a registered persona (digital identity) to every service he/she is using. The IP of the SP gives the user the identifiers and the credentials to access the SP. The next section describes the benefits and pitfalls of this identity management model.

4.1.1 Benefits and Pitfalls

What most characterizes this model in terms of benefits and pitfalls are the pitfalls it has. This model is most recognizable by the several access credentials users need to keep. Within the context of the traditional identity management model users have one digital identity for every service they use. Users cannot use the digital identity they have in one system in order to login into another. Services providers do not allow it. The traditional model implements the traditional password based method for authenticating users. This authentication method has many security drawbacks, but its simplicity is the primary reason for its wide use [25]. Having multiple access credentials to manage, leads users to be sloppier about it, as well as reluctant at the moment of adopting (other) online commerce solutions, which somewhat limits the scalability, the cost efficiency of the services and results in several of them not reaching their full potential [48-50]. This plurality of access credentials does not offer users full control over the disclosure and dissemination of their personal information to others [26-27]. Less frequently used credentials are easily forgotten and may even get in the possession of others very easily. Since users' personal information is scattered across several systems and there is no means by which users can monitor disclosed information, their privacy and anonymity may not be guaranteed. Moreover, the concerns (http://identityfight.org) of the whereabouts of their identity information, lead users to provide false information most of the times [51]. The services by themselves do not implement any mechanisms to certify that the information users provide is true and accurate. This leads users to be less accountable about the actions they take, because the users' real identity may not be discerned, making both users and service provides not reliable at all.

From the service providers' perspective, they have the double effort of providing the service, managing the users' information and make the authorization assertions in order to allow users to access its services. It leads to the duplication of the cost for the maintenance of the whole system. Nonetheless, the fact that within the context of this model the service provider deals the users' access to its services by its own identity management system, has the benefit of the information about the users being always firsthand information. Its integrity is therefore assured. Services using this line of work for managing their users' digital identity rarely implement secure communications, becoming more fragile to several kinds of attacks. This model for managing users' information is very sensible to the so called phishing attacks. These attacks are getting precise every day and because of it many services were offline and consequently, the costs for maintenance of the service rise.

In spite of this all flaws that this digital identity management model has, it is widely diffused, as it can be realized by our experience as users and by the number of credentials each one of us has to manage.

4.1.2 Password managers

Password managers are programs that concentrate all the digital identity credentials data into a single repository. This allows users to have a means by which they can manage their digital identity credentials in a more efficient and secure manner, avoiding typing the credentials data in the login forms. However, this kind of applications is still not avoiding the multi credentialism of users and, in order to use it, some kind of appropriated software is mandatory. Users still have to register at the services in the traditional way, by filling information forms. Users still may lose the access to any of their identities outside of the software scope. Thereby, it has the same flaws than the traditional digital identity management model since it is just a piece of software for users managing their credentials and not either a solution to strength users' privacy, anonymity and information scattering control, or a digital identity model in the first place. Plus, it gathers the credentials users have in one single place which increases the negative effect of a successfully identity deft attack. Even though it is, in one hand, a value for identity theft prevention, in another, it brings no relevant value to the protection of the privacy and anonymity of the users whatsoever (see lastpass.com).

4.2 Toward a unified login approach

In order to alleviate the users' credentials (identity) management overload, the services started to request an email address for the credentials (digital identity identifier) of every digital identity that access the services. This way, users only have to manage the different passwords they have for every digital identity they hold, when systems request a different and specific password format. This technique decreases the probability of users losing their access to services and eases the recovery of the access to their digital identity (account).

Using the users' email address for the representation of their digital identities does not protect both their privacy and anonymity. Otherwise, it increases the risks of a stolen identity and the probability of users loose either their privacy or (and) anonymity is now much higher, because the email address allows the discernment of users by the correlation of these (unique) identifiers. Moreover, users are still required to provide their particular information by filling forms. All the flaws of the traditional identity management model still existing in this approach.

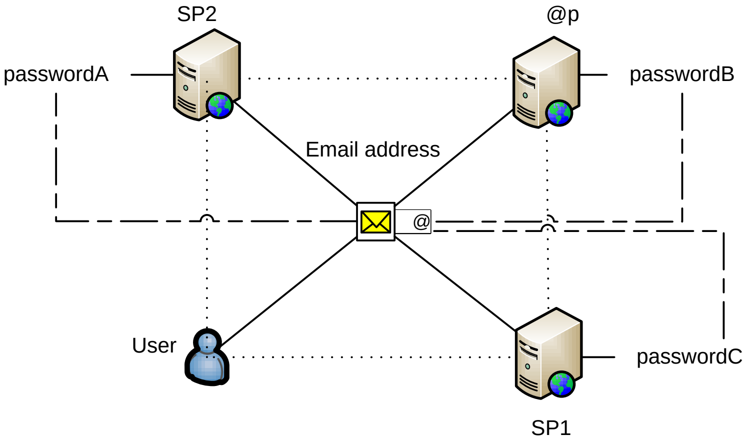


Figure 8 - Towards a user centric login approach.

In the Figure 8 it is illustrated a scheme of the association of the user with two service providers within a context of a user-centric login approach. The user has an email address provided by the @p email provider that is used as the credential username to access the SP1 and SP2 service providers. This way the user only has to manage the different passwords in each system, when they request different specific password formats.

4.3 The Federated Identity Model

The federated identity concept was introduced to address user digital identity management overload by creating a framework for systems interoperability. It establishes a set of rules and standards in order to create a reliable and secure namespace for the exchanging of identity information between source (usually users via their identity provider) and target (usually web applications), allowing users to keep their personal information less scattered across the Web applications, control its disclosure and dissemination, monitor its scrutiny and displacement, control their privacy, keep their anonymity protected, and access several web applications seamlessly by using only one digital identity [48].

The Federated Identity Model is a recent concept to address the management of digital identities. It is characterized for its user centric basis, focusing directly on the seamless user cross-domain. The cross-domain is achieved by the Single Sign On mechanism which improves the operational efficiency of all services within the federation domain. In order to enable users to cross domains seamlessly a set of agreements, standards and technologies have to be enable by the federated members. Within a federated identity context the identity provider and the service provider are detached [48]. This way the service providers only have to direct their attention to the service provisioning while the identity provider takes care of providing digital identities to users and managing their information. The identity provider is therefore the responsible agent for sending authorization assertions to service providers in order to enable users to access it.

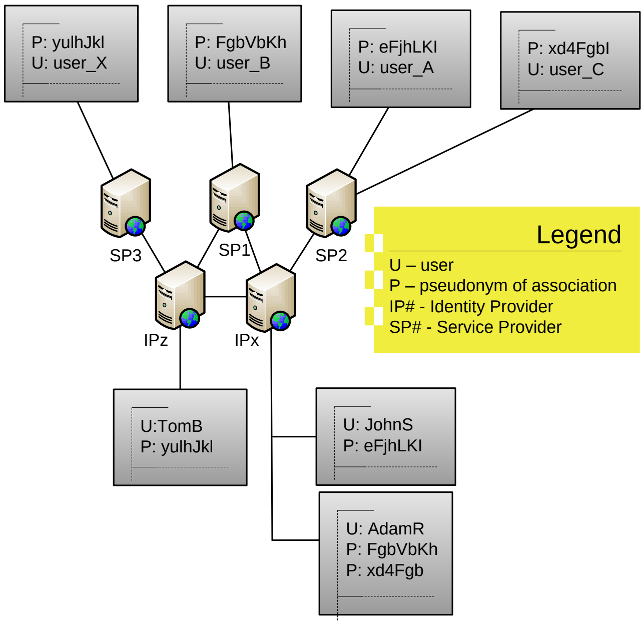


Figure 9 - Federated Identity Management Model Framework.

As it can be observed, the digital identity provisioning is done by a system detached from the ones that are providing the services (SPs). Within this federated context, users have only one set of credentials to manage in order to access all the services of the enabled federation. User "TomB" accesses the SP3 via "IPz", as it can be correlated by the pseudonym "yulhjkl". The user "JohnS" has a similar scenario as "TomB". The user "AdamR" uses two services, the SP1 and SP2, correlated by his pseudonyms "FgbVbKh" and "xd4Fb". The use of pseudonyms has the benefit of avoiding the discernment of users by the correlation of all the users' identifiers (pseudonyms).

The fact that the service providers are accepting assertions from the identity provider does not mean that some sort of circle of trust exist among them. The kind of trust required between a service provider and an identity provider depends highly on the type of assertions being exchanged, and the context in which they are being exchanged. It has more sense within the context of a company a high trust requisite to exist. Thereby, it makes sense if some sort of a company's service provider only accepts assertions from the identity provider inside the company (intranet), while for an online business provider it would make more sense to accept assertions from any identity provider in order to maximize the number of clients [23].

4.3.1 Benefits and Pitfalls

The Federated Identity Management Model entails significant benefits to all parties involved. The identity provider can focus on the improvement of the users' information management, such as users can have a better control over the disclosure and dissemination of their information and monitor its scrutiny and displacement as well. The identity provider can also focus on the improvement of the authentications methods (strong and mutual) as well as creating mechanisms to verify whether the information users provide is accurate or not. With the design of this model, users enjoy the benefits that single sign on provides them, since they only need one set of credentials to log in at the identity provider to access the services of the federation, seamlessly. This benefit comes together with the benefit of users not revealing their credentials to any of the service providers they access to, which contributes to the preservation of their privacy and anonymity. The redundancy and scalability problems of the older digital identity management systems are thereby decreased. The service providers delegate user account management tasks, such as password resets or profile information changes to the identity providers, while being enabled to acquire updated user information from them. The fact that service providers are relieved of the management of the users' information and authentication enables them to focus in the improvement of the business, which makes the federated identity model a leverage to boost the e-commerce Online [52].

Like any of the other models referred previously, the federated identity model is not free from drawbacks. The insertion of the single sign on mechanism requires exchanging of users' information between the identity provider and the service providers. Therefore, it demands that all parties secure their communication channels (implementation of HTTPS) against eavesdropping based attacks, in order to avoid undesirable leaks of information, but such demand is not always applied [23]. Within the context of this model the cost of a stolen credential is now much higher. If the attacker has in his/her possession the user credentials, the attacker gains access to all service providers the user is federated with and to all the user's personal information. The same is applied when the identity provider is ill-intended, but users may avoid such systems easily by checking the identity information provided by their browser and avoiding following doubtful information or adhering to suspicious systems.

4.4 Discussion

In this chapter two different models for the management of the users' identities and their access to the web applications (services) were introduced and described. Each model benefits and pitfalls were also analyzed. By the analyses effectuated to both identity management models it may be asserted that the most efficient model for managing the users' personal information and their access to the services, is the federated identity management model, either from users' point of view or from the point of view of the services providers. The federated identity model splits very well the jobs and concerns of each actor, by establishing a solution to regulate more efficiently the functionalities of each one. This model also reduces the impact of the drawbacks that could not be eliminated and that could be a potential prone of identity violation. Such is the case of the implementation of strong authentication mechanisms for avoiding the deep disadvantage of users loosing or be stolen the access to their federated digital identity. It must be also underlined that even though the federated identity model specifies guidelines for securing the users' identity data, users are the first verification and decision point in what regards to the certification of the security and reliability of the system (web application, service provider) which they desire to adhere to. A system that may not look suspicious could be an ill-intended identity provider or service provider and therefore users should avoid them. Therefore users should always certify whether the system is reliable by looking for references about it and/or check any available identity information that the browser they are using to access the service provides. In my opinion, in order to avoid ill intended identity or service providers that could put at stake the users' identity data, some sort of classification mechanism could be created by the most recognized identity or service providers, such that would supply users with a trust reference classification in order to let them know which system may be suspicious and which may not. This way the most diffused services Online would without doubt contribute to take the ill-intended identity and service providers out of business.

Due to its excellent benefits, federated identity should be implemented by the web applications instead of the traditional identity model. Nevertheless, its implementation may be slow and not so well articulated by the systems. The cost of modifying the existing digital identity management systems in order to implement federated identity is a tough hindrance to its implementation and diffusion across the web. However, such may be easily overcome by the potential of the federated identity model to leverage the boost of e-commerce transactions, and enrich-business market [52].

By the analysis of these models and approaches on digital identity management, it can be asserted that the digital identity management discipline has been getting closer to the user, making him/her the center of the management process and the decision maker in the management of the user’s identity. It may be said that the federated identity management model is the best model for managing users’ identity because it is more secure and reliable than the traditional model as it may be concluded by the analysis made to each model's benefits and pitfalls.

5 Federated Identity

The federated identity is outlined by a variety of particularities. These particularities should be known and understood in order to comprehend the federated identity concept. This chapter mentions and describes the most important characteristics of identity federation.

5.1 Federation Requirements

In order to establish a federation, the members must agree in a common of conditions [50]:

- Define a common framework built on industry standards, independent of specific implementations and networks as well as how its use will be accomplished.
- Assure the privacy safety of the users' information by keeping it secret, within the federation, and according to international privacy regulations. The users' personal information is only provided upon their approval.
- Protect the channels of information exchanging from any kind of attacks by applying to the security guidelines of the standards used for implementing identity federation.
- Assure the users have seamless access to any services of the federation and that their identity is only federated with any of the services of the federation upon their approval.
- Provide a way to establish trust amongst federation participants by setting a circle of trust between the members of the federations.
- Define rules for federation engagement, such as membership guidelines.
- Contribute for the common wealth of the federation.

5.1.1 Circle of trust

By setting a circle of trust the federated members agree to trust on:

- Any service within the federation.
- The Information from any service within the federation.

- The federated identity provider assertions.

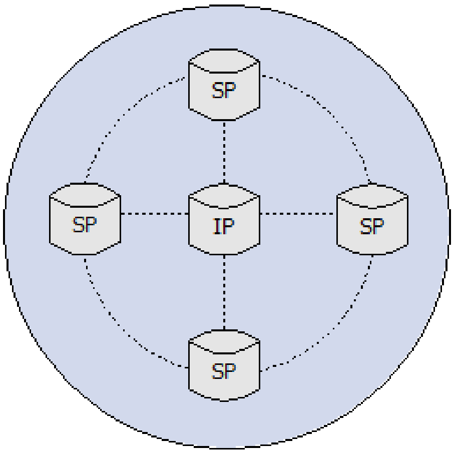


Figure 10 - Circle of Trust of an Identity Federation.

5.2 Who can be an identity provider?

In my point of view, when having trustworthy as a baseline of work, the most suitable entities to be an identity provider are:

- Governments.
- Banks.
- Telecommunications Enterprises.

Governments are the most suitable entity to be an identity provider, since they already manage the identities of their citizens. Plus, they already have the identity management structure defined and built. Moreover, most of the governments of the world already have their citizen identity information being managed by electronic systems, what could lead to an easy implementation of digital identity provisioning to Internet users.

Banks are other entities than are suitable to be identity providers for web users. They already have gathered accurate personal information of the clients and already have been providing home banking services to their clients. They could provide digital identity to their clients easily.

Telecommunications companies have the upside of owning a telecommunications network framework. Since they have accurate personal information about their clients, digital identity provisioning could be quickly implemented. Furthermore, they already provide digital identity to their clients from the fact that each one has a phone identity (phone number). By providing digital identity to their clients it would enable them to improve their business online which, somewhat would easy the commercial transactions.

5.3 Message Exchange Mechanisms

In order to enable identity federation between identity providers and service providers, identity messages must be exchange. Next will be mentioned the message exchanging mechanisms [25].

5.3.1 Browser-based

In a browser-based message exchange mechanism, the browser is the agent responsible for the bidirectional redirection of the requests between the service provider and the identity provider.

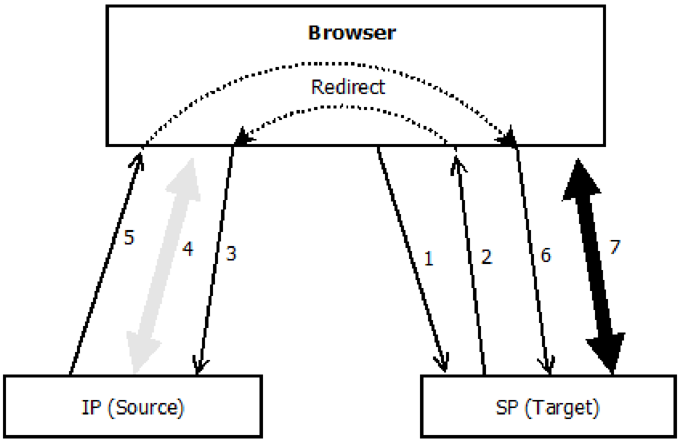


Figure 11 - Browser based message exchange.

Figure 11 illustrates the browser based message exchange flow when establishing a federation between an identity provider (IP) and a service provider (SP).

Assuming that no prior interaction existed before between any of the actors (user, SP and IP):

1. The user requests for a resource on the service provider.
2. The service provider not recognizing the user, builds an authentication request and sends it to the IP (chosen by the user), in this case, via HTTP redirect or HTTP post.
3. The IP receives the authentication requests, processes it and applies any relevant policy.

4. The IP authenticates the user.
5. The IP builds the authentication response and sends it to the SP via a browser redirect or a post.
6. The SP receives the authentication response from the IP and processes it.
7. The SP creates the necessary authentication context for the user (session) and redirects the user to the source request on 1.

5.3.2 Client-based

In a client based message exchange scenario the bidirectional redirection of the requests between the service provider and identity provider is established by a client installed on the users machine.

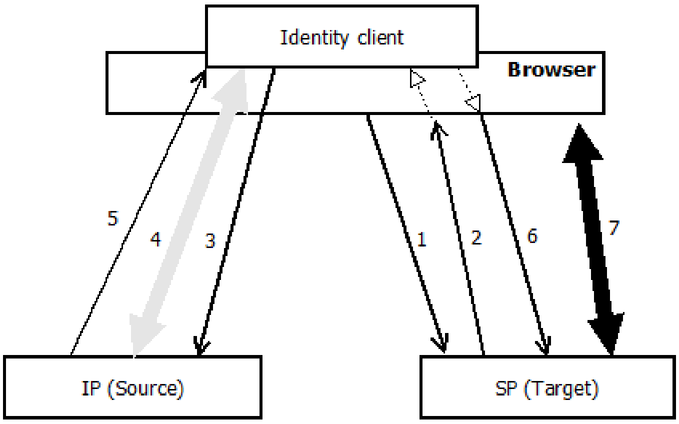


Figure 12 - Client based message exchange.

Assuming that no prior interaction existed before between any of the actors (user, SP and IP):

1. The user requests for a resource on the service provider.
2. The service provider not recognizing the user, builds a authentication requests and sends it to the IP (chosen by the user), in this case, in a special form understandable by the identity client. The browser recognizes the format and redirects the request to the identity client application.
3. The identity client processes the request and selects the identity provider, which may be indicated by the user.
4. The IP authenticates the user and builds the authentication response to send to the identity client.

5. The identity client processes the response, extracts the security token and may optionally cache it for later use. The identity client creates an authentication response with the secure security token and sends it to the SP.
6. The SP processes the authentication response received and creates an authentication context (session with the user's browser) for the user with the security token carried by the authentication response. The establishment of a session with the user's browser is made to avoid significant overhead in future re-authentications.
7. The SP redirects the user to the resource requested on 1.

When the user is already authenticated on the IP (point 4.) prior to a SP authentication request, he/she could be request for re-authentication by the SP authentication request. This is an option SPs have when requesting IPs to authenticate users.

5.4 Account linking

When an identity provider is asserting an identity to a service provider, the linking of the users' identity between the identity provider and the service provider may be established in two different ways.

5.4.1 Direct Linking

The direct linking of a user's identity is made by the use of a unique identifier. The Figure 13 illustrates it.

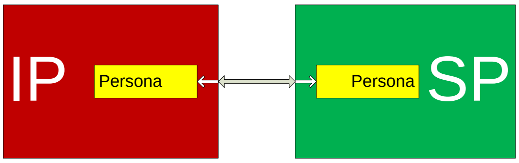


Figure 13 - Direct linking.

The direct linking of the user's identity (Persona) with service providers raises privacy issues for the user. Since global unique identifiers are being used for the linking of the user's identity between the IP and SP, the user's identity as well as their general common activities can be easily discerned by the service providers from the correlation of their identity identifiers. Figure 14 illustrates one possible scenario of identifiers correlation.

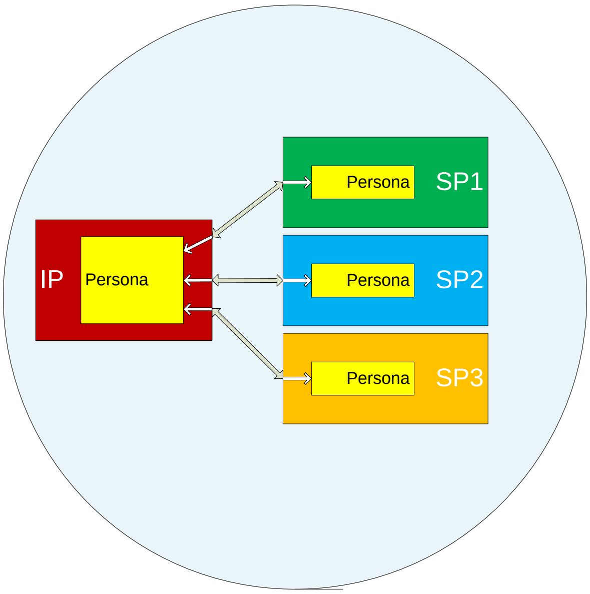


Figure 14 - Federated direct linking.

The Figure 14 illustrates that if the SPs correlate their identity data information, they will be enabled to discern the user easily.

5.4.2 Indirect linking

In the indirect linking of the users' identity between the IP and SP, the identity provider asserts a per-unique (web application) identifier (pseudonym) to the users' identity for the representation of the user in the service provider as illustrated in the Figure 15.

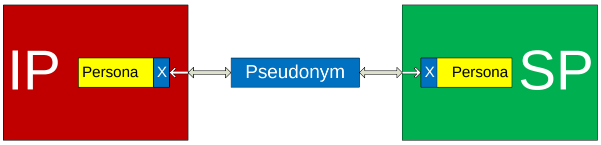


Figure 15 - Indirect linking.

This way the identity of the user cannot be discerned by the correlation of the pseudonym, protecting the users' privacy as it is illustrated in Figure 16.

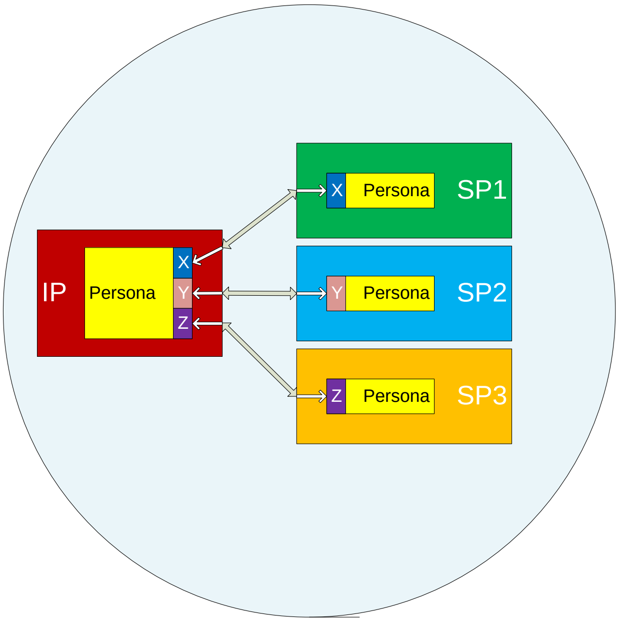


Figure 16 - Federated Indirect Linking.

If the service providers try to correlate their users' identifiers they won't realized that the X, Y and Z identifiers actually represent the same user in all services. This way it is proved that pseudonymity leads to anonymity.

6 Federated Identity State of Art

This section is addressed to the existing initiatives for the implementation of federated identity management systems. The main ones are mentioned, described, explained and analyzed in order to find out through its benefits, pitfalls and the need of users which, standards are more suitable for the development of the Global identity federated identity provider, the identity provider to be developed under the study result of this master thesis. The standards will not be deeply explained from the technical point of view, but rather from their main features, because it would be redundant to make a thorough analysis of every protocol, since right now the differences between most of them are minimal, and they either follow the same guidelines or converged into a unified line of work. In different standards, there are similarities in their main features as well as in its implementation and resources used. Therefore, since the principles behind each standard converge in most of their specifics, the aim of the analysis taken is to find out which federated identity standard(s) could have a wide range of interoperability and acceptance by the online systems and which one(s) is/are more suitable to implement on the GlobaliD identity provider. Discontinued federated identity standards will not be taken in consideration.

Upon the selection of the initiative(s) for implementing federated identity in the GlobaliD identity framework, the chosen one(s) will be explained in a deeper detail in a section for the purpose. The analysis of the state of the art will start from the point that there is no single industry standard meeting all of the federation requirements. The following table indicates the most known existing standards that approach the federated identity management model. Table 2 lists the standards that will be discussed:

Table 2 - Federated Identity Initiatives.

Initiative	Consortium	Version	Date	Type
OpenID	The OpenID Foundation	2.0	12/2007	Browser
Windows Live ID	Microsoft	6.0	7/2007	Browser
Liberty ID-FF	Liberty Alliance	1.2	11/2003	Browser Client
WS-Federation	IBM, Microsoft,VerySign	1.1	12/2006	Browser Client
Shibboleth	Internet2	1.3		Browser
Identity Metasystem	Microsoft	1.0	7/2007	Client
SAML	OASIS	2.0	03/2005	Browser

In section 4.3, the federated identity model was analyzed within the context of online services. Therefore in order to address to the state of art of the existing standards for implementing identity federation, I will start with the most diffused ones online, the OpenID and the Windows Live ID.

6.1 OpenID

OpenID [53] was originally developed by Brad Fitzpatrick in the summer of 2005, to authenticate commentators in LiveJournal’s blogs in order to avoid spam. The OpenID is a simple, open, decentralized, free framework, user-centric Internet digital identity system, that is highly distributed and diffused in/by the many services online [54]. The purpose of the OpenID is to provide users with single sign-on in services that do not require strong security. It has its own authentication protocol [55], such roughly follows the browser-based mechanism described previously. However, it deviates in first steps, adding communication between target site and source site prior to the main redirect sequence.

The OpenID protocol takes advantage of already existing internet technology such as URI, HTTP, SSL, Diffie-Hellman. It was designed from the beginning to use only standard HTTP(S) requests and responses, not requiring any special additional capabilities of the user-agent or other client software. OpenID start from the point that users already have digital identities online such as blogs, photo sharing, etc. Firstly, it used the URI or XRI as globally unique identifiers for the representation of each user identity. But in the version

2.0 of the protocol it provided a non-mandatory pseudonym mechanism for avoiding the discerning of users by the correlation of URIs in order to protect their anonymity. The OpenID last version also provides mechanisms for the implementation of strong authentication, such as the use of smartcards, but this factor is optional.

URI's were initially chosen as identifiers because they were already and widely adopted identifiers, which were not only unique but also easily recognizable. This way a user could easily tie content to his identity by using the URI of his personal web page as an identifier, which was particularly convenient in blogging environments were the user's identifier would be the URL of his own blog. The use of URI as identifiers can easily dereference and lead to the users' identity provider as well as it immediately exposes user information to any party that comes across the identifier. XRI is an additional resolution mechanism for URIs by OASIS. Therefore they were adopted as suitable identifiers for OpenID and were introduced in the OpenID 2.0 specification.

OpenID does not mandate any specific structure of personae and/or accounts. However it is obviously expected by OpendID designers that the primary way of OpenID usage is the use of single or few globally-unique identifiers. The OpenID has the flexibility of users being capable to delegate the authentication of their URI to any identity provider, so that if one deems that the identity provider is not trustworthy anymore or is going to end services in an foreseeable future, he/she may simply switch to another identity provider, without losing his/her URI identity reference.

OpenID is not free of flaws, it has some security limitations [56]. Besides the pitfalls that it has for being an identity federation mechanism, certain indications must be followed in order to avoid eavesdropping and man-in-the-middle attacks. However these indications are only optional. OpenID users are still forced to manage duplicate information in multiple services and if the service requires any extra information about the user, it must collect that from the user, validate it, if necessary, and store it locally. OpenID agents cannot interoperate with agents using other identity protocols.

6.2 Windows Live ID

Windows Live ID [57] (WLID) is the last evolutionary state of the Microsoft Internet Passport. It enables users with a decentralized authentication feature. For the representation of users, the windows live id uses per-site specific unique identifiers (unique pseudonyms)

[58]. Windows live id allows an association of all the WLIDs the users have, this way they only need to login once at the WLID’s IP to access the services associated with the others WLID’s of the association. Any other information that the web applications request about the users must be provided by themselves by typing it in the forms the services provide for it or by authorizing the IP to allow the services to access the information. This last feature is only performed with the user prior authorization. The Windows Live ID may be accessed using WS-Trust, CardSpace or even Authentication Dial In User Service (RADIUS). It may be used in an identity federated scenario by using WS-Federation [59], as well. Moreover, the WLID became an OpenID IP recently.

6.3 The Liberty Alliance Project

The Liberty Alliance Project [60] is an initiative from the Liberty Alliance. The Liberty Alliance is a consortium of more than 150 organizations including technology vendors, consume-facing companies, educational organizations and governments from around the world, as well as hundreds of additional organizations that participates in Liberty’s various open community Special Interests Group (SiGs). The aim of the Liberty Alliance Project is to provide open standards, guidelines and best practices for network identity management, within a federation context. Decentralization and openness are the main goals of the Alliance. In order to address that The Alliance released some specifications [61] in three cumulative phases:

- **Phase1 –** Identity Federation Framework (ID-FF) - ID-FF is the core of the framework, federated identity services, single sign-on, session management, account linking, and privacy. This phase is now merged with SAML2.0.
- **Phase2** – Identity Web services Federation (ID-WSF) - ID-WSF enables the users’ management sharing of their attributes and ID to SP and personalized Services.
- **Phase3** – Identity Services Interface Specifications (ID-SIS) - services providing such as contact book, calendar, geo-location, presence, profiling.

All the three phases are designed to build on top of each other: Phase 2 (ID-WSF) relies on the Phase 1 (ID-FF) and Phase 3 (ID-SIS) instantiates the ID-WSF framework.

6.3.1 ID-FF

The Identity Federation Framework (ID-FF) is the core of the Liberty Federation framework. The ID-FF follows both browser and client based message exchange as described in section 5.3. ID-FF supports mechanisms for authentication context, e.g. the user of smart cards. The ID-FF 1.2 specifications were used by OASIS to be converged with their SAML V1.1 specification and the Internet2 Shibboleth, forming the foundation of SAML V2.0. The differences between the version 2.0 of SAML and the ID-FF specifications are of structural and formatting matter and not change any of the characteristics that ID-FF already had. Therefore I will not provide a promiscuous description of the ID-FF 1.2 characteristics. Later I will introduce SAML and its features. Moreover, SAML V2.0 substituted the ID-FF specification in the alliance project architecture.

6.3.2 ID-WSF

The ID-WSF builds upon ID-FF to provide a framework for identity-based web services in a federated network identity environment. It specifies a SOAP-based invocation framework, defining a SOAP binding message for the user information exchanging between identity providers and service providers that can be established over the HTTP protocol. The ID-WSF does not specify any contents for the SOAP Body, allowing the development of identity services within the context of ID-WSF. The body of the SOAP messages can be used to share the users' personal information across identity and service providers by keeping their security and privacy untouched. The WS-Security from WS-* is the specifications that provides this functionality to ID-WSF.

6.3.3 ID-SIS

ID-SIS is a collection of interoperable identity services. It is built on top of the ID-WSF framework, which permits web services to communicate among them by using SOAP over HTTP calls. The specifications define the specific syntax and semantics for sharing different slices of identity attributes over the services. The services provided by the ID-SIS can include ones such as registration, contact book, calendar, geo-location, presence, or alerts.

The next figure illustrates the architecture of the Liberty Alliance Federation Framework.

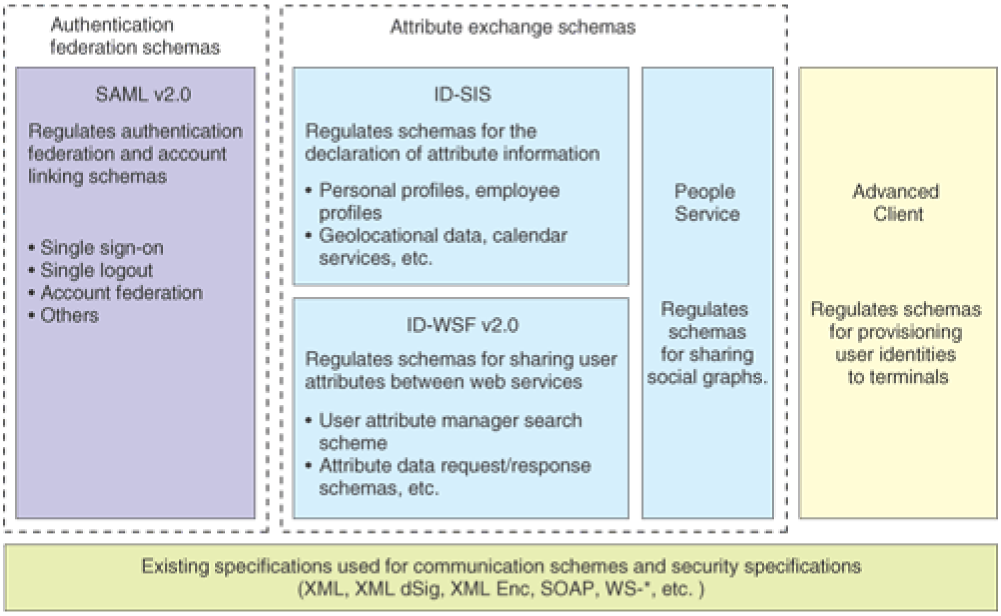

Authentication federation schemas
Attribute exchange schemas
SAML v2.0
Regulates authentication federation and account linking schemas
• Single sign-on
• Single logout
• Account federation
• Others
ID-SIS
Regulates schemas for the declaration of attribute information
• Personal profiles, employee profiles
• Geolocational data, calendar services, etc.
ID-WSF v2.0
Regulates schemas for sharing user attributes between web services
• User attribute manager search scheme
• Attribute data request/response schemas, etc.
People Service
Regulates schemas for sharing social graphs.
Advanced Client
Regulates schemas for provisioning user identities to terminals
Existing specifications used for communication schemes and security specifications (XML, XML dSig, XML Enc, SOAP, WS-*, etc.)

Figure 17 - The Liberty Alliance specifications architecture.

6.4 WS-* Services

The WS-Federation Services [62] is an initiative of many companies such as Microsoft, IBM, BEA, IBM, Microsoft, RSA Security and VeriSign. It relies on other models in order to enable identity federation among services. The WS-Federation is a group of specifications used together in order to activate identity federation to entities. The WS-Federation Services are defined by a group of specifications:

- WS-Security.
- WS-Secure Conversation.
- WS-Policy.
- WS-Federation.
- WS-Trust.

6.4.1 WS-Security

The WS-Security specifications define the SOAP security extensions providing data integrity and confidentiality. It defines how to attach extensions and encryption to SOAP messages, and the insertion of the XML security tokes in the headers of the messages. WS-

Security does not support secure communications, it relies on the WS-SecureConversation to achieve the security of the messages exchanging channel.

6.4.2 WS-SecureConversation

The WS-SecureConversation defines the rules by which security tokens are created and shared between the evolving parties of the federation. It uses the security context tokens to apply it. The WS-SecureConversation defines how the process of the encrypting and signing keyshare computed by the parties.

6.4.3 WS-Policy

The WS-Policy defines the rules to express the capabilities and requirements of entities used in Web services environments by the use of policy assertions, in conjunction with the WS-Security SOAP extensions.

6.4.4 WS-Federation

WS-Federation extends WS-Trust to provide flexible federated identity architecture with clean separation between trust mechanisms, security token formats, and the protocol for obtaining tokens. It uses the browser based mechanism for the entities interaction. In conjunction with WS-Security, WS-Trust and WS-Policy, it provides the support for secure propagation of identity, attribute authentication and authorization information. It also enables the brokering of trust and security token exchange as well as support for privacy by hiding identity and attributes information. It provides federated sign-out, as well. The WS-Federation leave the privacy decisions to implementers what may lead to bad practices, increasing the risk of privacy violation.

6.4.5 WS-Trust

WS-Trust is the initiator of the circle of trust within the federation. It defines the type of security tokens they must use in the headers of WS-Security. WS-* does not provided support for authentication context (e.g. use of smart card).

6.5 Shibboleth

The Shibboleth standard [63] is a specification by Internet2 that is now converged with the SAML specifications. Shibboleth standard follows the browser-based model for the messages exchanges between entities. This specifications address federation services within an institutional context rather than within a web applications one. It does not mandate the model of a common IP for the federation but the acceptation of the credentials of the others services providers' identity providers by the federated members (credentials recognizance). Further, it defines the standard-based protocol for securely transferring attributes between home site and resource site. These traits demands that Shibboleth adds optional WAYF (Where Are You From) service for identity provider selection besides the elements identity provider and service provider. Shibboleth provides single sign-on to federation and attributes exchange built on SAML Specifications. The specifications define Shibboleth-specific transient identifiers, but recommends that it should be used only once. Shibboleth specifies control over login methods, by defining authentication context indications.

6.6 Microsoft Information Cards and the Microsoft identity metasystem vision

The identity metasystem is a vision of interoperability between existing digital identity management specifications, with the idea that a universal identity system is unlikely to ever exist [64]. It is based upon a set of principles called the “Laws of Identity” [21] described in subsection 2.10. The identity metasystem is implemented by the Card Space technology, also called information cards, as they are the base of the identity metasystem. These cards have a list of claims from the user and a list of assertion types the identity provider supports. In order to achieve interoperability between system supporting a different federated identity specification, claim transformers are used for crossing organizational boundaries.

In the identity metasystem the identity provisioning is achieved by the identity selector mechanism. Users are the ones that choose which identity (information cards) to provide from a set they have on their card space. The idea behind this is to rely on the users the assertions of digital identities to service providers as they already do in the offline world, by choosing which card to provide in any given situation. Each card possesses a visual representation with a card-like picture and a name describing the underlying identity. Is not demanded that each card possesses any particular information of the user, it can have only information about the identity provider that manages the holder identity. The technology enables a given user to customize the information he/she wants to provider for the card. These cards have a great advantage of being strong phishing resistant, what protects the privacy of the user and the intervenient on the information cards exchange from ill-intended attacks. Since an information card is given in any authentication method, password based authentication method is avoided.

The identity metasystem is built by using the Web Services specifications (WS-*) [62] described before. The messages are secured using WS-Security, the encapsulating protocol used for claims transformation is WS-Trust and the format and claims negotiations between participants are conducted using WS-MetadataExchange and WS-SecurityPolicy (which is based on WS- Policy). The identity metasystem follows a client based message exchange mechanism between identity provider and service provider.

The information cards can be of the following kinds:

- Managed Cards – are given by the identity provider to the users, who previously imported it to the identity selector.
- Personal Cards or self-issue cards – are associated to self-identity provisioning, they are stored in the user machine.

The CardSpace implements a mechanism to support (strong) authentication context, it implements the personal pin identifier (PIN).

The card space idea assumes that the users are always in the possession of their information cards, which may lead to a situation that a user cannot access some website due to not having his/her CardSpace at the moment. The following table shows the list of available claims in the information card technology.

Table 3 - Identity Claims of the Identity metasystem.

Claim	**URI**
Given Name	http://schemas.xmlsoap.org/ws/2005/05/identity/givenname
Last Name	http://schemas.xmlsoap.org/ws/2005/05/identity/surname
Street	http://schemas.xmlsoap.org/ws/2005/05/identity/streetaddress
Locality (City)	http://schemas.xmlsoap.org/ws/2005/05/identity/locality

State or Province	http://schemas.xmlsoap.org/ws/2005/05/identity/stateorprovince
Postal Code	http://schemas.xmlsoap.org/ws/2005/05/identity/postalcode
Country/Region	http://schemas.xmlsoap.org/ws/2005/05/identity/country
Phone Number	http://schemas.xmlsoap.org/ws/2005/05/identity/homephone
Other Phone	http://schemas.xmlsoap.org/ws/2005/05/identity/otherphone
Mobile Phone	http://schemas.xmlsoap.org/ws/2005/05/identity/mobilephone
Date of Birth	http://schemas.xmlsoap.org/ws/2005/05/identity/dateofbirth
Gender	http://schemas.xmlsoap.org/ws/2005/05/identity/gender
PPID	http://schemas.xmlsoap.org/ws/2005/05/identity/privatepersonalidentifier

6.7 SAML 2.0

SAML [65] stands for Security Assertion Markup Language. It is an OASIS [66] specification that defines a XML-based framework for communicating security and identity information between computing identities. The last version of SAML, version 2.0, was a convergence of SAML 1.1, Shibboleth 1.3 and Liberty ID-FF 1.1 specifications. SAML workflow baseline is defined by assertion, protocols, bindings and profiles. Such is structured in three main parts:

- SAML Core [67] (assertions and protocols) – defines the xml format for the messages exchanged between federated agents and specify the use of it.
- SAML Bindings [68] – specify a set of bindings to ensure that independently implemented SAML-conforming software can interoperate when using standard messaging or communication protocols by defining the mapping of the SAML messages exchanged.
- SAML Profiles [69] – outlines a set of rules describing the body of the messages for carrying the assertions, and how they are extracted from and embedded into a transporting protocol.

SAML V2.0 standard define other specification to approach other federated identity requirements.

- SAML Metadata [70] – defines the syntax for entities bootstrap the trust process when enabling federation for a user.

SAML also defines guidelines for defining authentication context [71], and released a set of considerations regarding privacy and security [72], as well as a set of requirements for its implementation [73].

6.8 Discussion

The last chapter provided a description of the features of the most relevant federated identity model standards.

The OpenID initiative is addressed to web applications that do not require strong security. It is more addressed to SSOn scenarios rather than identity data management. The Windows Live ID only provides specifications for the service provider side, does not define any authentication context. Both previous initiatives have the advantage of being wide diffused across the Web.

The identity metasystem has the advantage of trying to create a platform of translation between the several existing standards than rather creating a new one. Even though it is a good idea, it must implement mechanisms of translation addressed to each standard that it tries to integrate, becoming the specification too complex. It also has the drawback of being necessary to install appropriated software in the users systems and by the implementation of the Windows CardSpace it assumes that the users always have it with them. Nevertheless, the concept of information card is excellent for structuring the users' personal information.

The WS-* set of specifications is a complex initiative. It leaves the implementation of important features such as pseudonyms to the developers, rather than imposing it. It may lead to a situation where two WS-* Federation system may not be able to federate due to different implementation. It does not implement any authentication context.

Shibboleth is more addressed for corporations and institutional contexts and together with Liberty ID-FF converged on SAML 1.1, resulting on SAML 2.0 specifications.

The SAML specifications define the protocol messages format and the bindings for the embedding and transport of protocol messages between systems. It also defines a set of profiles for structuring the information exchange according to the actors involved, providing a base line for the use of SAML assertions and protocols to accomplish specific use cases or achieve interoperability when using SAML features. SAML specifies how systems define the authentication context required for the establishment of identity federation and security lines to secure the information exchange, as well.

SAML demands the use of pseudonyms for the representation of users, avoiding the possibility of services discern users by the correlation of their identity identifiers. The use of pseudonyms for the representation of users protects the users' privacy and so on their anonymity. Upon the previous descriptions and analyzes and when comparing SAML with the other initiatives, it may be realized that SAML is the most complete and suitable standard to integrate in the GlobaliD Federated Digital Identity framework Actually, since every other standard either uses SAML on their specifications or converged into it, SAML is the *de-facto* initiative for implementing identity federation.

Now that the standard to use in the development of GlobaliD Federated Digital Identity is chosen, its particularities will be explained in a more deep detail, in the following section.

7 SAML 2.0

SAML defines a group of specifications that addresses a wide range of aspects for implementing identity federation. The specifications address the format of the messages exchanged between the service provider and identity provider on the assertions and protocols specifications. In another specification it addresses the bindings for exchanging the protocol messages. In the profiles specifications it describes the coordinated use of the protocols messages and the bindings for establishing message exchange between systems. SAML also addresses a specification for system configuration data exchange in the Metadata specifications. Other specifications are defined to approach other identity federation particularities such as the authentication context establishment.

In a SAML session the following actors may be involved:

- Requester – the entity who creates the authentication request and to whom the correspondent response is to be returned.
- Presenter – the entities who presents the request.
- Requested Subject – the entity about whom one or more assertions are being requested.
- Relying Party – the entity or entities expected to consume the assertion to accomplish a purpose defined by the profile or context of use, generally to establish a security context.
- Identity Provider – the entity to whom the presenter gives the request and from whom the presenter receives the response.
- The principal – a reference to the user.

Next it will me mentioned and described the particularities of SAML in nutshell based on the SAML specifications available on www.oasis-open.org [67-73] and according to the features used for the development of the GlobaliD identity provider. The specifications used on the GlobaliD identity provider will be described more in detail. The reading of the following subsections does not exempt the reader of this document from consulting the specifications provided on the website referred previously in order to comprehend the SAML initiative.

7.1 Assertions

The core specifications are made by the assertions and protocols definitions [67]. They specify the syntax and semantics for the assertions messages about authentication, attributes, and authorization, and for the protocols that convey this information. SAML assertions and protocol messages are encoded in XML and use XML namespaces to refer to it. They are typically embedded in other structures for transport, such as HTTP POST (binding) request or XML-encoded SOAP messages. An assertion is a package of information that supplies zero or more statements made by a SAML authority about a subject. SAML authorities are sometimes referred to as assertion parties in discussions of assertion generator and exchange, and systems entities that use received assertions are known as relying parties. Usually, assertion parties and relying parties are called identity providers and services providers, respectively. The SAML specification defines three different kinds of assertion statements that can be created by a SAML authority, which are associated to a given subject. The three kinds of statement defined in this specification are:

- **Authentication** – the assertion subject was authenticated by a particulars means at a particular time.
- **Attributes** – the assertion subject was associated with the supplied attributes.
- **Authorization decision** – a request to allow the assertion subject to access the specified resource has been granted, denied or indeterminate.

The outer structure of an assertion is generic, providing information that is common to all of the statements within it. Within an assertion, a series of inner elements describe the authentication, attribute, authorization decision or user-defined statements containing the specifics. Some elements are of option use and others are either required or mandatory.

The <Assertion> XML element specifies the basic information that is common to all assertions, including the following elements and attributes:

- Version [Required] - specifies the SAML current version used for building the respective assertion.
- ID [Required] – holds the unique identifier for the assertion.
- Issue Instant [Required] – holds the time instant of the issue of the assertion in UTC.

- <Issuer> [Required] – defines the SAML authority that is making the claim(s) assertion and provides information about it. The issue should be ambiguous to the intended relying parties.
- <ds:Signature> [Optional] – holds the XML signature that protects the integrity of and authenticates the issuer of the assertion.
- <Subject> [Optional] – defines the subject statement in the assertion, i.e., the subject of all of the statements in the assertion. The subject of the assertion may be encrypted by using a <EncryptedID> element.
- <Conditions> [Optional] – specifies the conditions that must be evaluated when assessing the validity of and/or when using the assertion. Defines constraints on the acceptable use of SAML assertions, such as the validity period of the assertions, the audience, and its retention for future use.
- <Advice> [Optional] – contains any additional information that the SAML authority wishes to provide to the relying party.
- <Statement> [Optional] – it is an extension point that allows other assertion-based applications to reuse the SAML assertion framework. This holds any of the following statements:
  - <AuthnStatement> – this element describes that the assertion subject was authenticated by a particular means at a particular time.
  - <AuthzDecisionStatement> – describes a statement by the SAML authority asserting that a request for access by the assertion subject to the specified resource has resulted in the specified authorization decision on the basis of some optionally specified evidence. It states whether the subject of the assertion may access the determined resource for take a determined action, e.g., the permission to read some file.
  - <AttributeStatement> – describes a statement by the SAML authority asserting that the assertion subject is associated with the specified attributes. The specified attributes may be encrypted. In such a case the element <EncryptedAttribute> is used instead of the <Attribute> one.

Multiple statements may be provided to the assertion or none at all. When used at least one authentication statement in the assertion the <Subject> of the assertion must be provided as well. In order to allow the use of other kinds of assertions and statements not defined by the specifications, SAML permits the extensions of the schemas in order to support it.

Next there is an example of a SAML assertion message:

```
<saml:Assertion Version="2.0" ID="4c0129fc-0e05-496b-a9a0-00977bc7a973"
IssueInstant="2010-05-30T00:27:51Z" xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion">
  <saml:Issuer>www.globalid.com</saml:Issuer>
  <saml:Subject>
    <saml:NameID Format="urn:oasis:names:tc:SAML:2.0:nameid-format:persistent">
     45311d3b-d019-4de2-9d04-4f60a6077069
</saml:NameID>
    <saml:SubjectConfirmation Method="urn:oasis:names:tc:SAML:2.0:cm:bearer">
<saml:SubjectConfirmationData NotOnOrAfter="2010-05-30T00:57:51Z"
Recipient=”www.sp.com/assertionconsumerservice.aspx”
InResponseTo="7a08de47-5f80-42a6-895b-444ab55c642a" />
    </saml:SubjectConfirmation>
  </saml:Subject>
  <saml:Conditions NotOnOrAfter="2010-05-30T00:57:51Z" />
  <saml:AuthnStatement AuthnInstant="2010-05-30T00:27:51Z">
    <saml:AuthnContext>
      <saml:AuthnContextClassRef>urn:oasis:names:tc:SAML:2.0:ac:classes:SmartcardPKI
      </saml:AuthnContextClassRef>
    </saml:AuthnContext>
  </saml:AuthnStatement>
  <saml:AttributeStatement>
    <saml:Attribute Name="Mobile"
                    NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:basic"
FriendlyName="Personal">
      <saml:AttributeValue xsi:type="xs:string" xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-nstance">960000000
</saml:AttributeValue>
</saml:Attribute>
 <saml:Attribute Name="Email" NameFormat="urn:oasis:names:tc:SAML:2.0:attrname-format:basic"
FriendlyName="Work">
<saml:AttributeValue xsi:type="xs:string" xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">userwork@email.com</saml:AttributeValue>
    </saml:Attribute>
    <saml:EncryptedAttribute>
      <EncryptedData Type="http://www.w3.org/2001/04/xmlenc#Element"
xmlns="http://www.w3.org/2001/04/xmlenc#">
        <EncryptionMethod Algorithm="http://www.w3.org/2001/04/xmlenc#aes256-cbc" />
        <KeyInfo xmlns="http://www.w3.org/2000/09/xmldsig#">
          <EncryptedKey xmlns="http://www.w3.org/2001/04/xmlenc#">
            <EncryptionMethod Algorithm="http://www.w3.org/2001/04/xmlenc#rsa-1_5" />
            <KeyInfo xmlns="http://www.w3.org/2000/09/xmldsig#">
              <X509Data>
<X509Certificate>MIIBozCCAQygAwIBAgIES/p9yTANBgkqhkiG9w0BAQUFADAWMRQwEgYDVQQDEwt3d3cuaWRwLmNvbTAeFw
0xMDA1MjQxMzIzMjFaFw0yMDA1MjExMzIzMjFaMBYxFDASBgNVBAMTC3d3dy5pZHAu
Y29tMIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCcKZ6i7MFmK9R+TfEpgL6gmZba7oqL4K
Bh0uS2RPYcnRjKeCfa0ZIn/7hKw78g8MymIXtp3EflXvgGCeyrjai8biTh1INnHev9lIN+gwnS
Y1+nWwp7nPR4WxYTR6bFBdRlWff+h6EiAlRjcgbQ13iXTh1W9do+M78Jzlokw3ku+wIDAQABMA
0GCSqGSIb3DQEBBQUAA4GBACDC8UPfZNDWjiXMJIgvPYCWTl7D/KhRVFPQk9CdYGS78aOqUbhJ
Vv7RSigWnmsurgfu6Gy5n+4VQBL0b8JteqnnoRQOEvK6mQsAcO8Ug+lQLr1GoQqpmBrccQjSzQ
P8Mw5JDV84OgurKger9a6FDoAmbUXiX9R5HN4EaQ0hdwEP
                </X509Certificate>
              </X509Data>
            </KeyInfo>
            <CipherData>
              <CipherValue>
OiUzH56fTHT5vToJq6EWTeFztrx+TBXSKNPTrgsv0u/gyh3VbMkE8tNpze49P2kR5wKOJmhCaje2iJ9UX9Cealnf6ZCe8qY4kNS
hbUPOtQcAv9eHbLcG8tQH0ss7kmmTxbQMwJOu/5shfKfxMcfBSi+E+HMrWDjKoVgnJRkRL+c=</CipherValue>
            </CipherData>
          </EncryptedKey>
        </KeyInfo>
        <CipherData>
          <CipherValue>
d4HAtr18hCGnHRGbXkK8n7MptdTgTP75KCCT8DJ0wHuxlhsRYbxI6ieUQi4bbCTe3WFqala2SZqtadpdEhqbM83y6QGLllvLeph
v1Q+he8j8QYyn2dK41kHt54P1l18od1/OQq62V9FrvNhgh4lYQLBE4s98HvDP91o/mkOgmlKkcunAhHxZ6BLV3a67Xv3UM/TC7C
```

```
TWrAr4kk2GWMM0FPlgbAd9yRAQh3W5SpZN2U/AFiPRCt73MrUFgBCfqXHrO6DUeaNZ+kUA+rkEj1Nta9Mak9QbacqoByi6dLvTg
xS4Q+ptLMu8O8zu3s4dXJe8GDM0iD1ZNl8pSEONZuREwxfQluMF2ldmkDU5ZyDwtA37seS8RDdmr+AHc3rKAYnxjfCQHYcmC6Dk
tFAbgk1lO0CoBpVr2s8Ae9rdDwZukDIfxxaUZQFQOcP6ENIJyleA1lP7OtNR3yEe1WNc2pFp6GjRide48zvpqHjnNv0t3Rv46F2
Q5RGlu4na5Ufyf//T</CipherValue>
        </CipherData>
      </EncryptedData>
    </saml:EncryptedAttribute>
  </saml:AttributeStatement>
</saml:Assertion>
```

In this SAML assertion it can be seen all the XML nodes defined before. The assertion has the ID: "4c0129fc-0e05-496b-a9a0-00977bc7a973"; issued at: "2010-05-30T00:27:51Z". It may be seen the issuer of the assertion is: www.globalid.com. The SAML NameID holds the pseudonym asserted to the user with the value of: 45311d3b-d019-4de2-9d04-4f60a6077069. The subject confirmation information and conditions are also indicated on the subject confirmation and conditions XML element. The user is authenticated at the IP by the use of a smart card with PKI capabilities. The assertion provides the user mobile and email in the attributes XML element. The assertion also provides an encrypted attribute. The public key certificate of the identity provider is supplied the keyinfo XML element of the assertion. In the end of the assertion it is provided the signature of the message.

7.1.1 SAML Identifiers

SAML defines a series of identifiers for specifying the format of the identifier asserted to the user (<Subject>). SAML demands that some of these identifiers must be pseudonyms in order to protect the identities privacy and anonymity.

Some of the defined SAML identifiers are:

- Unspecified – leaves to individual implementation the interpretation of the content of the element.
- Email Address – specify that the user identifier is an email address.
- Entity Identifier – specify the issuer of a SAML message.
- Persistent Identifier – indicates that the identifier is a persistent opaque identifier for a principal that is specific to an identity provider and a service provider. Persistent name identifiers generated by identity providers must be constructed using pseudo-random values.

- Transient Identifier – indicates that the content of the element is an identifier with transient semantics and should be treated as an opaque and temporary value by the relying party. Transient Identifiers should be constructed using pseudo-random values.

SAML defines other types of identifiers on the assertion and protocols specification section 8. Refer to it for further insights on this matter.

7.2 Protocols

SAML Protocols messages can be generated and exchanged using a variety of protocols. The SAML bindings specifications [68] describes specific means of transporting protocol messages using existing widely deployed transport protocols. The protocols specifications define two base types of messages: Request messages and Response Messages. Figure 18 illustrates the process of the messages by the authorities and relying parties.

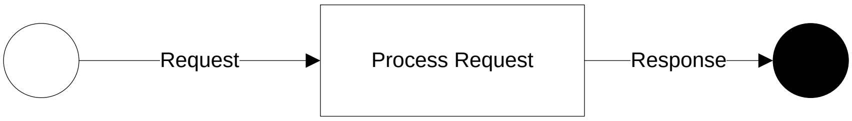


Figure 18 - SAML Request-Response Protocol.

In certain cases, when permitted by the SAML profile being used, a SAML response may be generated and sent without the responder having received a corresponding request.

The protocols defined by SAML achieve the following actions:

- Returning one or more requested assertions. This can occur in response to either a direct request for specific assertions or a query for assertions that meet particular criteria.
- Performing authentication on request and returning the corresponding assertion.
- Registering a name identifier or terminating a name registration on request.
- Retrieving a protocol message that has been requested by means of an artifact.
- Performing a near-simultaneous logout of a collection of related sessions (“single logout”) on request.
- Providing a name identifier mapping on request.

7.2.1 Requests and Responses

SAML Requests are messages sent by the relying party to the asserting party. SAML Responses are messages sent by the asserting party to the relying party.

All SAML messages have or may have the following elements and attributes:

- Version [Required] - specifies the SAML current version used for building the respective assertion.
- ID [Required] – holds the unique identifier for this assertion.
- Issue Instant [Required] – the time instant of the issue of the assertion in UTC.
- Destination [Required] – an URI reference indicating the address to which this request has been sent. This is useful to prevent malicious forwarding of requests to unintended recipients, a protection that is required by the protocols bindings.
- Consent [Optional] – indicates whether or not (and under what conditions) consent has been obtained from a principal in the sending of this request.
- <saml:Issuer> [Required] – defines the SAML authority that is making the claim(s) assertion and provides information about it. The issue should be ambiguous to the intended relying parties.
- <ds:Signature> [Optional] – hold the XML signature that protects the integrity of and authenticates the issuer of the assertion.

  <Extensions> [Optional] – to allow custom scheme agreed on between the communicating parties.

In the case of a SAML response the following elements or attributes may apply:

- InResponseTo [Optional] – holds a reference to the identifier (ID) of the request to which the response corresponds, if any.
- <Status> [Required] – specifies a code representing the status of the corresponding request in its inner element <StatusCode>. It may return a message to the operator in the inner element <StatusMessage> of <Status>. The <StatusCode> may provide a subordinated second level status code to provide more specific reasons for the failure of the authentication, e.g., a status code of a response may have a second-level status code of NoAuthnContext,

meaning that the authentication context required by the requester was not obtained by the requester.

Next there is an example of a SAML Response message:

```
<samlp:Response ID="1ec77a4a-a89f-4679-abd9-2afa942c0d20" InResponseTo="f49b20ca-385a-4d3f-bc36-
dfed91e03d6b"
                Version="2.0" IssueInstant="2010-05-30T01:06:23Z"
Destination="www.sp.com/assertionconsumerservice.aspx"
                xmlns:samlp="urn:oasis:names:tc:SAML:2.0:protocol">
  <saml:Issuer Format="urn:oasis:names:tc:SAML:2.0:nameid-format:entity"
xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion">
    www.ip.com/ssoservice.aspx</saml:Issuer>
  <samlp:Status>
    <samlp:StatusCode Value="urn:oasis:names:tc:SAML:2.0:status:Success" />
    <samlp:StatusMessage>User authenticated successfully</samlp:StatusMessage>
  </samlp:Status>
  <saml:Assertion Version="2.0" ID="999deccb-619a-426e-ba80-8042e5a4a4fb" IssueInstant="2010-05-
30T01:06:23Z"
                 ... shortened not to spend too much space ...
  </saml:Assertion></samlp:Response>
```

In this SAML response it can be seen all the XML elements and nodes defined before. The same response specifies its Id on the ID attribute of <samlp:Response> element. Other information characteristic of a SAML response is also provided, such as the version and the Issue Instant. The InResponseTo attribute holds the authentication request id of authentication request message that originated this assertion. In the response it may be also seen the description of the issuer of the response and the status code of the response stating that the user was successfully authenticated at the identity provider. Then one assertion is provided to define the user for whom the response was created.

7.2.2 Authentication Request Protocol

The authentication request protocol defines the XML messages for requesting assertions containing authentication statements to establish a security context at one or more relying parties. In such case a <AuthnRequest> message element is sent to a SAML authority and it replies with a <Response> message containing one or more such assertions. In order to request that an identity provider issue an assertion with an authentication statement, a presenter sends it an <AuthnRequest> message that describes the properties that the resulting assertions needs to have to satisfy the purpose. The <AuthnRequest> message may have the following elements or attributes:

- <NameIDPolicy> [Optional] – specifies constrains on the name identifier to be used to represent the requested subject. This element defines whether the identity provider is allowed to create the subject identifier and respective type.
- <saml:Conditions> [Optional] – specifies the SAML conditions the requester expects to limit the validity and/or use of the resulting assertion(s).
- <RequestedAuthnContext> [Optional] – specifies the requirements, if any, that the requester places on the authentication context that applies to the responding provider authentication of the presenter.
- ForceAuthn [Optional] – holds a Boolean value. If “true”, the requester requires that the identity provider authenticates the presenter directly rather than rely on a previous security context.
- IsPassive [Optional] – holds a Boolean value. If “true”, the requester requires that the identity provider do not take visible control of the user interface, but rather interact with the presenter in a noticeable fashion.
- AssertionConsumerServiceURL [Optional] – specifies by value the location to which the <Response> message must be returned to the requester.
- ProtocolBinding [Optional] – defines a URI reference that identifies a SAML protocol binding to be used when returning the <response> message.
- ProviderName [Optional] – specifies the human-readable name of the requester for use by the presenter’s user agent or the identity provider.

Next there is an example of a SAML request message:

```
<samlp:AuthnRequest ID="b3a8066d-53e1-4cee-b668-de8f6446bd70" Version="2.0"
IssueInstant="2010-05-30T01:33:21Z" Destination="www.globalid.com/ssoservice.aspx"
Consent="urn:oasis:names:tc:SAML:2.0:consent:current-implicit" ForceAuthn="true"
IsPassive="false" ProtocolBinding="urn:oasis:names:tc:SAML:2.0:bindings:HTTP-POST"
AssertionConsumerServiceURL="www.sp.com/assertionconsumerservices.aspx"
                    xmlns:samlp="urn:oasis:names:tc:SAML:2.0:protocol">
  <saml:Issuer xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion">www.sp.com
</saml:Issuer>
  <Signature xmlns="http://www.w3.org/2000/09/xmldsig#">
    <SignedInfo>
      <CanonicalizationMethod Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#" />
      <SignatureMethod Algorithm="http://www.w3.org/2000/09/xmldsig#rsa-sha1" />
      <Reference URI="#b3a8066d-53e1-4cee-b668-de8f6446bd70">
        <Transforms>
<Transform Algorithm="http://www.w3.org/2000/09/xmldsig#enveloped-signature" />
          <Transform Algorithm="http://www.w3.org/2001/10/xml-exc-c14n#">
<InclusiveNamespaces PrefixList="#default samlp saml ds xs xsi"
xmlns="http://www.w3.org/2001/10/xml-exc-c14n#" />
          </Transform>
        </Transforms>
        <DigestMethod Algorithm="http://www.w3.org/2000/09/xmldsig#sha1" />
        <DigestValue>oe/Lgi2Wvc1uuhVH24eB2VVdca8=</DigestValue>
      </Reference>
    </SignedInfo>
```

```
<SignatureValue>e9E50T8abWyLwYPVi2WZu7xmHSUJc4jwuWkyvzpcvjhnKyX79b/9cbxDtHcA/QJ2fZrkm7yqv2/dHUBt+A8
RcXK0ud1QAGUqfvipdBVgsUEV451efsHSAyzD1LYfa67o/9Ow2a0RKYyPhIU7zdyR3s8NL9Db7GqUHzp6YnwlHhE=
  </SignatureValue>
    <KeyInfo>
      <X509Data>
        <X509Certificate>
MIIBozCCAQygAwIBAgIES/p9yTANBgkqhkiG9w0BAQUFADAWMRQwEgYDVQQDEwt3d3cuaWRwLmNvbTAeFw0xMDA1MjQxMzIzMjF
aFw0yMDA1MjExMzIzMjFaMBYxFDASBgNVBAMTC3d3dy5pZHAuY29tMIGfMA0GCSqGSIb3DQEBAQUAA4GNADCBiQKBgQCcKZ6i7M
FmK9R+TfEpgL6gmZba7oqL4KBh0uS2RPYcnRjKeCfa0ZIn/7hKw78g8MymIXtp3EflXvgGCeyrjai8biTh1INnHev9lIN+gwnSY
1+nWwp7nPR4WxYTR6bFBdRlWff+h6EiAlRjcgbQ13iXTh1W9do+M78Jzlokw3ku+wIDAQABMA0GCSqGSIb3DQEBBQUAA4GBACDC
8UPfZNDWjiXMJIgvPYCWTl7D/KhRVFPQk9CdYGS78aOqUbhJVv7RSigWnmsurgfu6Gy5n+4VQBL0b8JteqnnoRQOEvK6mQsAcO8
Ug+lQLr1GoQqpmBrccQjSzQP8Mw5JDV84OgurKger9a6FDoAmbUXiX9R5HN4EaQ0hdwEP
        </X509Certificate>
      </X509Data>
    </KeyInfo>
  </Signature>
<samlp:NameIDPolicy Format="urn:oasis:names:tc:SAML:2.0:nameid-format:persistent"
AllowCreate="true" />
<samlp:RequestedAuthnContext>
<saml:AuthnContextClassRef
xmlns:saml="urn:oasis:names:tc:SAML:2.0:assertion">urn:oasis:names:tc:SAML:2.0:ac:classes:Smartcard
PKI</saml:AuthnContextClassRef>
  </samlp:RequestedAuthnContext>
</samlp:AuthnRequest>
```

In this SAML response it can be seen all the XML elements and nodes defined before. The authentication request has the ID of ="b3a8066d-53e1-4cee-b668-de8f6446bd70" and was issued at the referred issue instant. The destination of the message is the single sign on service of the globalid.com identity provider. It is also defined that the SAML binding used for sending the authentication message was the HTTP POST binding. In the assertion consumer service attribute it is provided the address to which the response provided by the replier of the authentication request should be sent to. The other elements specify the signature and the public key certificate of the service provider, the authentication request issuer.

Yet, SAML defines other protocols for approaching the same authentication process. These are going to be briefly introduced. For further insights please refer to the SAML specifications.

SAML defines the Artifact Protocol, a specialized protocol for providing a mechanism by which SAML protocol messages can be transported in a SAML binding by reference instead of by value. It also defines a protocol for changing the subject identifier after the identity provider had established one. The Name Identifier Mapping Protocol is defined for providing the identity provider a means by which it may obtain a name identifier for the same subject in particular format or federation namespace. The specifications provide a SAML-defined Identifiers collection for referring to common access actions, subject name identifier formats and attribute name formats, as well.

Finally, SAML defines a Single Logout protocol for providing a message exchange protocol by which all sessions provided by a particular session authority are near-simultaneously terminated.

7.3 Bindings

In order to provide a specification for the exchanging of the protocol messages, SAML defines a series of bindings [68] for the purpose. On the bindings specifications it is defined how the SAML messages (request-response) are mapped into the transports protocol in order to implement interoperability between federated members. In this section it will be mentioned the bindings that SAML defines for the SAML agents establish message exchanging. The binding implemented on the GlobaliD identity provider will be described more in detail. For deep insights on the available SAML bindings please refer to the SAML Bindings specifications.

SAML provides the following protocol bindings:

- SOAP binding – a lightweight protocol intended for exchanging structured information in a decentralized, distributed environment. The two major design goals for SOAP are simplicity and extensibility. It defines how to use SOAP to send and receive SAML requests and responses, normally carried on the SOAP body of the XML message envelope defined by SOAP.
- Reverse SOAP (PAOS) binding – is a mechanism by which an HTTP request can advertise the ability to act as a SOAP responder or a SOAP intermediary to a SAML requester.
- HTTP Redirect binding – defines a mechanism by which SAML protocol messages can be transmitted within URL parameters. Permissible URL length is theoretically infinite, but unpredictably limited in practice. Therefore, specialized encodings are needed to carry XML messages on a URL, and larger or more complex message content can be sent using the HTTP POST or Artifact bindings. This binding may be composed with the HTTP POST binding and the HTTP Artifact binding to transmit request and response messages in a single protocol exchange using two different bindings.
- HTTP Post binding – defines a mechanism to transport SAML protocol messages within the base64-encoded content of an HTML form control. This

binding may be used with either the HTTP Redirect binding or the HTTP Artifact binding. This binging was the chosen for implanting on the identity provider developed for this document. It will be explained more in detailed further in this section.

- HTTP Artifact binding – is a specialized binding for the transmission of SAML messages by reference using a SAML stand-ion called an artifact. A separate, synchronous binding, such as the SAML SOAP binding, is used to exchange the artifact for the actual protocol message using the artifact resolution protocol defined in the SAML assertions and protocols specification. This artifact-based binding is the most secure binding that the specifications define.
- URI binding – URIs are protocol-independent means of referring to a resource. Thus, this binding use the URIs to carry messages containing the assertion.

Some considerations must be undertaken when using any of the available SAML bindings. Some bindings define a “RelayState” mechanism for preserving and conveying state information, such as the resource the user tried to access before the authentication request took place. This relay state must be preserved during the subsequent message exchanging in order to enable the service provider to redirect the user to the resource the user tried to access in the first place. The protocol bindings should use of SSL 3.0 or TLS 1.0, servers must authenticate to clients using X.509 v3 certificate. Authentication of both the SAML requester and the SAML responder associated with a message should be established according the context in use.

7.3.1 HTTP POST Binding

The HTTP Post binding defines a mechanism by which SAML protocol messages may be transmitted within the base64-encoded content of an HTML form control. This binding may be composed with the HTTP Redirect binding and the HTTP Artifact binding to transmit request and response messages in a single protocol exchanging using two different bindings. The HTTP Post binding is intended for cases in which the SAML requester and responder need to communicate using an HTTP user agent as an intermediary, such as a web browser. This may be necessary, for example, if the communicating parties do not share a direct path of communication. It may also be needed if the responder requires an interaction with the user agent in order to fulfill the request, such as when the user agent must authenticate to it. Messages are encoded for use with this binding by encoding the XML into an HTML form control and are transmitted using the HTTP POST method. A SAML protocol message is form-encoded by applying the base-64 encoding rules to the XML representation of the message and placing the result in a hidden control within the HTML form. The form control may include a relay state data indicating, which service provider resource the user was trying to access to before the authentication request took place. Figure 19 illustrates the messages exchanging flow when using the HTTP Post binding for making either SAML Requests or SAML Response.

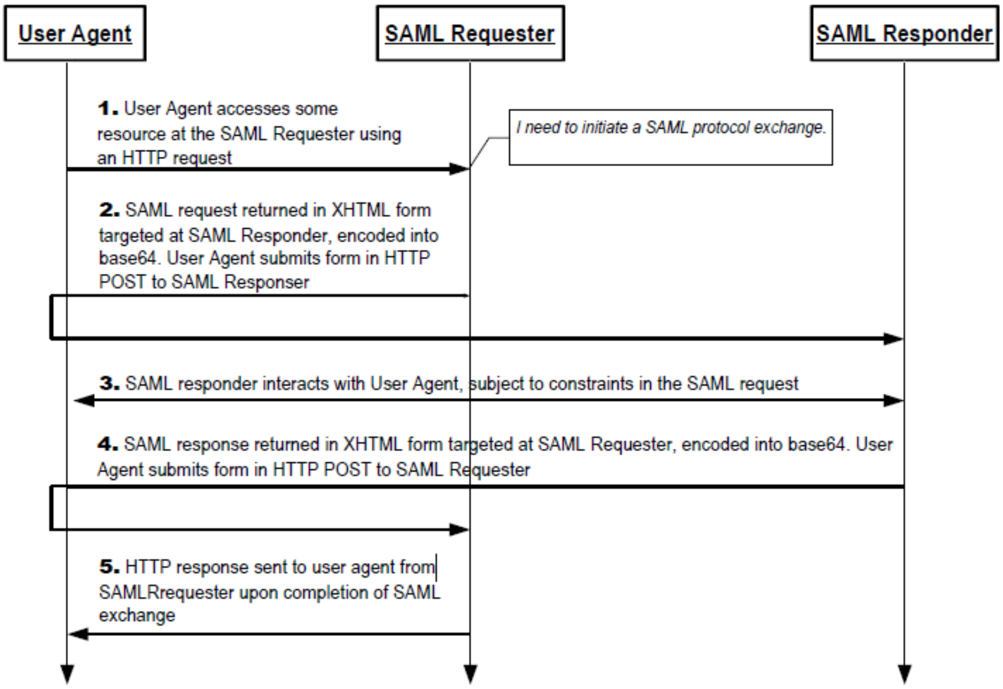

User Agent
SAML Requester
SAML Responder
1. User Agent accesses some resource at the SAML Requester using an HTTP request
I need to initiate a SAML protocol exchange.
2. SAML request returned in XHTML form targeted at SAML Responder, encoded into base64. User Agent submits form in HTTP POST to SAML Responser
3. SAML responder interacts with User Agent, subject to constraints in the SAML request
4. SAML response returned in XHTML form targeted at SAML Requester, encoded into base64. User Agent submits form in HTTP POST to SAML Requester
5. HTTP response sent to user agent from SAMLRrequester upon completion of SAML exchange

Figure 19 - HTTP POST binding message exchanging flow.

1. Initially, the user agent makes an arbitrary HTTP request to a system entity. In the course of processing the request, the system entity decides to initiate a SAML protocol exchange.
2. The system entity acting as a SAML requester responds to an HTTP request from the user agent by returning a SAML request. The request is returned in an XHTML document containing the form and content of the SAML message.
3. In general, the SAML responder may response to the SAML request by immediately returning a SAML response or it may return arbitrary content to facilitate subsequent interaction with the user agent necessary to fulfill the

request. Specific protocols and profiles may include mechanisms to indicate the requester's level of willingness to permit this kind of interaction (for example, the IsPassive attribute in <samlp:AuthnRequest>.

4. Eventually the responder should return a SAML response to the user agent to be returned to the SAML requester. The SAML response is returned in the same fashion as described for the SAML request in step 2.
5. Upon receiving the SAML response, the SAML requester returns an arbitrary HTTP response to the user agent.

7.4 Profiles

The profiles specifications [69] define the use of SAML assertions and request-response messages in communications protocols and framework, as well as profiles that define SAML attribute value syntax and naming conventions. One type of SAML profile outlines a set of rules describing how to embed SAML assertions into and extract them from a framework or protocol. Such profile describes how SAML assertions are embedded in or combined with other objects (for example, files of various types, or protocol data unites of communication protocols) by an originating party, communicated from the originating party to a receiving party, and subsequently processed at the destination.

The intent of this specification is to specify a selected set of profiles of various kinds in sufficient detail to ensure that independently implemented products will interoperate. A set of profiles is defined to support single sign-on (SSO) of browsers and other client devices. The set of profiles defined by the SAML profiles specification are:

- Enhanced Client or Proxy (ECP) Profile – an enhanced client or proxy (ECP) is a system entity that knows how to contact an appropriate identity provider possibly in a context-dependent fashion, and also supports the Reverse SOAP (PAOS) binding. This profile specifies interactions between enhanced clients or proxies and service providers and identity providers for enabling the authentication of users.
- Identity Provider Discovery Profile – in deployments having more than one identity provider, service providers need a means to discover which identity provider(s) a user uses. This discovery profile defines a profile by which a service provider can discover which identity providers a user is using with the

web browser SSO profile. This profile relies on a cookie that is written in a domain that is common between identity providers and service providers in a deployment. The domain that the deployment predetermines is known as the common domain in this profile, and the cookie containing the list of identity providers is known as the common domain cookie.

- Single Logout Profile – usually users authenticate at several websites while browsing the web. This profile defines a profile to enable users to single sign out of all sessions established with n multiple providers at nearly-simultaneous time. The profile allows the protocol to be combined with any of the defined SAML Bindings.
- Name Identifier Management Profile – this profile is used to implement scenarios in which, either the identity provider or the service provider wishes to change the subject identifier value or to inform that it will no longer accept or send messages using a particular identifier.
- Web Browser SSO Profile – relies on a scenario supported by the web browser, where a web user either accesses a resource at a service provider, or accesses an identity provider such that the service provider and desired resource are understood or implicit. The web user authenticates (or has already authenticated) to the identity provider, which then produces an authentication assertion (possibly with the input from the service provider) and the service provider consumes the assertion to establish a security context for the web user.

This last profile specifies the steps for message exchanging, which may vary depending on the binding used in each step. Figure 20 illustrates the basic template for achieving SSO.

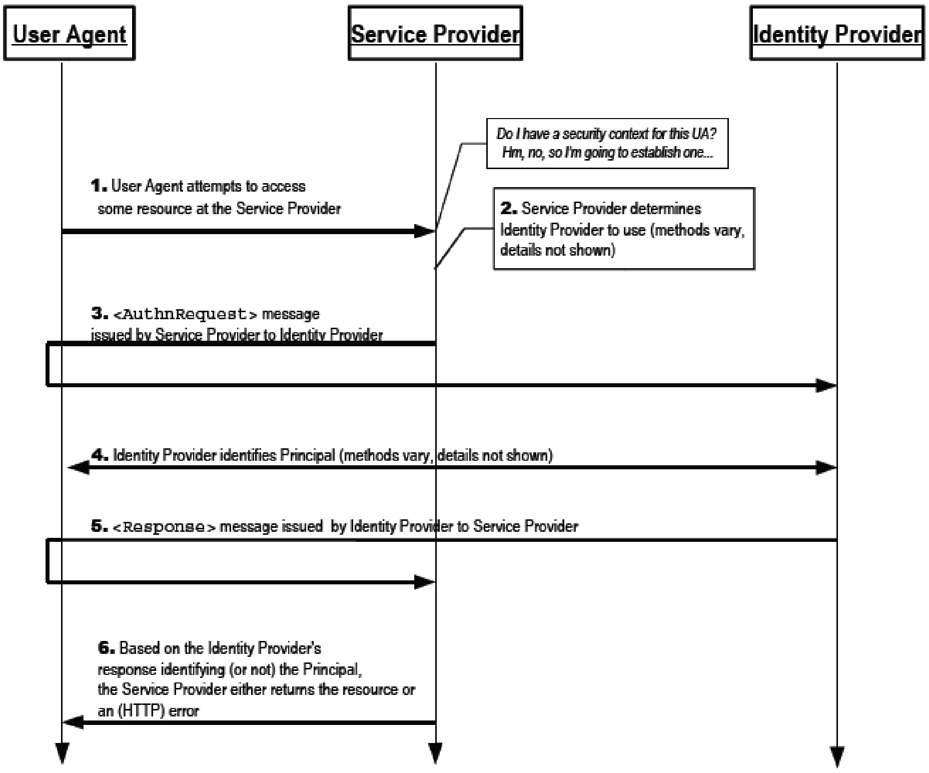

User Agent
Service Provider
Identity Provider
Do I have a security context for this UA? Hm, no, so I'm going to establish one...
1. User Agent attempts to access some resource at the Service Provider
2. Service Provider determines Identity Provider to use (methods vary, details not shown)
3. <AuthnRequest> message issued by Service Provider to Identity Provider
4. Identity Provider identifies Principal (methods vary, details not shown)
5. <Response> message issued by Identity Provider to Service Provider
6. Based on the Identity Provider's response identifying (or not) the Principal, the Service Provider either returns the resource or an (HTTP) error

Figure 20 - Web Browser SSO Profile message exchanging scheme.

1. HTTP Request to Service Provider – in step 1, the principal, via an HTTP User Agent, makes an HTTP request for a secured resource at the service provider without a security context.
2. Service Provider Determines Identity Provider – in step 2, the service provider obtains the location of an endpoint at an identity provider for the authentication request protocol that supports its preferred binding. The means by which this is accomplished is implementation-dependent. The service provider MAY use the SAML identity provider discovery profile.
3. <AuthnRequest> issued by Service Provider to Identity Provider – in step 3, the service provider issues an <AuthnRequest> message to be delivered by the user agent to the identity provider. The HTTP Redirect, HTTP POST, or HTTP Artifact binding can be used to transfer the message to the identity provider through the user agent.
4. Identity Provider identifies Principal – in step 4, the principal is identified by the identity provider by some means outside the scope of this profile. This may

require a new act of authentication, or it may reuse an existing authenticated session.

5. Identity Provider issues <Response> to Service Provider – in step 5, the identity provider issues a <Response> message to be delivered by the user agent to the service provider. Either the HTTP POST, or HTTP Artifact binding can be used to transfer the message to the service provider through the user agent. The message may indicate an error, or will include (at least) one authentication assertion. The HTTP Redirect binding must not be used, as the response will typically exceed the URL length permitted by most user agents.
6. Service Provider grants or denies access to Principal – in step 6, having received the response from the identity provider, the service provider can respond to the principal's user agent with its own error, or can establish its own security context for the principal and return the requested resource.

Note that an identity provider can initiate this profile at step 5 and issue a <Response> message to a service provider without the preceding steps.

SAML also defines profiles for non-browser dependency. They describe the use of the protocol with the same name with a specific synchronous binding such as the SOAP binding. The profiles are:

- Artifact Resolution Profile.
- Assertion Query/Request Profile.
- Name Identifier Mapping Profile.
- SAML Attribute Profile.

7.5 Authentication Context Specification

The authentication context specification [71] defines syntax for the definition of authentication context declarations and an initial list of authentication context classes. Each class defines a proper subset of the full set of authentication contexts. Classes have been chosen as representative of the current practices and technologies for authentication technologies, and provide asserting and relying parties convenient shorthand when referring to authentication context issues.

For instance, an authentication authority may include with the complete authentication context declaration it provides to a relying party an assertion that the authentication context also belongs to an authentication context class.

This intends to:

- Make it easier for the authentication authority and relying party to come to an agreement on what are acceptable authentication contexts by giving them a framework for discussion.
- Make it easier for relying parties to indicate their preferences when requesting a step-up authentication assertion from an authentication authority.
- Simplify for relying parties the burden of processing authentication context declarations by giving them the option of being satisfied by the associated class.
- Insulate relying parties from the impact of new authentication technologies.
- Make it easier for authentication authorities to publish their authentication capabilities, for example, through WSDL.

The authentication context declaration is defined by the use of an URI for reference it. The GlobaliD identity provider prototype will use the following two URIs for specifying which available authentication context the user is currently using:

- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:PasswordProtectedTransport
- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:SmartcardPKI

Other important URIs for the GlobaliD framework are:

- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:MobileOneFactorUnregistered
- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:MobileTwoFactorUnregistered
- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:MobileOneFactorContract
- URI: urn:oasis:names:tc:SAML:2.0:ac:classes:MobileTwoFactorContract

Other URIs are defined to specify a series of different authentication context that systems may implement. Please refer to the specifications for deep insights about it.

7.6 Metadata specifications

SAML profiles require agreements between system entities regarding several aspects for implementing identity federation, such as the definition of supported bindings, endpoints, certificates and keys, and so forth. The metadata specifications [70] define how the different systems establish trust among them, by defining protocols for assert the necessary configurations in order to make the deployment of SAML systems entities easier. The metadata is organized around an extensible collection of roles representing common combinations of SAML protocols and profiles supported by the system entities. Such roles included that of SSO Identity Provider, SSO Service Provider, Affiliation, Attribute Authority, Attribute Requester and Policy Decision Point. Each role is described by an element derived from the extensible base type of RoleDescriptor. Such descriptors are in turn collected in the <EntityDescriptor> container XML element, the primary unit of SAML metadata. In its metadata, the entities may alternatively represent an affiliation of other entities, such as an affiliation of service providers by the use of the <AffiliationDescriptor>.

The aims of the metadata specifications are to provide a variety of security mechanisms for the establishment of trustworthiness between SAML entities, particularly with the ability to individually sign most of the metadata messages, for enabling entities to authenticate among themselves.

SAML defines the <EndPointType> and <IndexedEndPointType> elements for entities define their available SAML protocol binding at which they can be sent SAML protocol messages. The <EndPointType> consists of the following attributes:

- Binding [Required] – the URI defining the binding supported.
- Location [Required] – the URI attribute that specifies the location of the endpoint.
- ResponseLocation [Optional] – optionally specifies a different location to which response messages sent as part of the protocol or profile should be sent.

The <IndexedEndPointType> extends the <EndPointType> with:

- Index [Required] – a required attribute that assigns a unique identifier value to the end point so that it can be referenced in a protocol message.
- IsDefault [Optional] – an optional additional Boolean attribute used to designed the endpoint by default.
- The <EntityDescriptor> element specifies metadata for a single SAML entity. It consists of the following elements:

- entityID [Required] – specifies the unique identifier of the SAML entity whose metadata is described by the element's content.
- ID [Optional] – a document-unique identifier for the element typically used as a reference point when signing.
- ValidUntil [Optional] – optional attribute indicates the expiration time of the metadata contained in the element and any contained elements.
- cacheDuration [Optional] – optional attribute indicates the maximum length of time a consumer should cache the metadata contained in the element and any contained elements.
- <ds:Signature> [Optional] – an XML signature that authenticates the containing element and its contents.

The metadata also provides the <IDPSSODescriptor> and <SPSSODescriptor> by which entities may specify some requirements and provide their endpoints. The <IPSSODescriptor> is used by identity providers to define requirements such as the requirement for the message received to be signed or the SSOService endpoint to which the authentication request can be sent. The <SPSSODescriptor> is used by the service provider to define requirements such as whether it wants the identity providers to send them signed assertions or not and to specify the assertion consumer service endpoint. SAML metadata defines other elements for specifying the responsible organization and people for the SAML entity in the <Organization> and <ContactPerson> elements. Metadata should be publishing in a well-known URL location or in a zone of their corresponding DNS.

SAML published the technical requirements and a set of documents that describes features that are mandatory and optional for implementations claiming conformance to SAML V2.0 on the SAML conformance specifications [73]. It also published a set of security and privacy considerations [72] for providing information to architects, implementers, and reviewers of SAML-based systems about the following:

- The privacy issues to be considered and how SAML architecture addresses these issues.
- The threats, and thus security risks, to which a SAML-based system is subject.
- The security risks the SAML architecture addresses, and how it does so.
- The security risks it does not address.

- Recommendations for countermeasures that mitigate those security risks.

8 GlobaliD Federated Identity Framework

As it was said in the introduction of this Master's thesis, personal information sharing is a very often online activity. Users provide their particular information either when registering at some web application or to any other user. Most often they publish several kinds of data, such as text, photos, voice, and video. The exchange means users use to share their personal information lacks the mechanisms to provide full control over the disclosure and dissemination of the personal information they share, usually do not offer information scrutiny and displacement monitoring or any mechanisms to certify that the information representing them is accurate or asserted as true or false. Most of the times users are somehow forced to give up about some privacy or even their anonymity in order to accomplish a task, such as share a photo, an email, etc… Very often they multiply the provisioning of their information or the information of others across many platforms without any kind of concerns about it. Usually, the channels of communications used to exchange their information are not secure and several kinds of attacks are easily succeeded. Therefore most implemented identity management systems of today, are not offering the reliable and secure mechanisms necessary to relieve users about any concerns they might have when sharing their particular information with others. Moreover, most of the identity management systems do not follow any security guide lines, such as, implementation of encryption protocols in order to prevent unauthorized access to the data [48].

The proposed GlobaliD framework aims to provide a model for digital identity management in order to supply users with an easy, comfortable and most of all secure means for the management of their particular information as well as efficient means for the services providers managing the users' access to their systems and retrieve required data about them. The GlobaliD intends to be a resemblance of the OpeniD but instead of using an own protocol for establishing identity federation, it uses the SAML specifications and implements particular own features.

In order to address this aim a GlobaliD federated identity provider prototype will be developed by using the proposed GlobaliD federated digital identity framework. With the development of the GlobaliD federated identity provider prototype it is intended to show that the use of the GlobaliD framework for federated identity management strengthens the users' privacy, anonymity, accountability, trustworthiness, accuracy and veracity of the personal information they share within a web context as well as the security of the communications channels used for exchanging the information. It also improves the adherence and the access to different service providers as well as the managing of the users' personal data.

The GlobaliD framework accomplishes this aim by:

- Providing the secure and seamlessly access to the web applications.
- Implementing strong authentication means for the secure access to the identity provider.
- Offering of a relatively simple interface that allows users to selectively choose the pieces of personal information they want to provide to web applications and others.
- Providing a mechanism which users can monitor their personal information scrutiny and displacement.
- Providing web application adherence and information sharing log.
- Providing a mechanism to assert about the accuracy of the information (true, false, managed).
- Providing a ranking mechanism to classify the digital identity and or personae over the veracity of the information.
- Securing the means of information exchanging, regarding the best practice guidelines to assure the integrity of the information, the privacy safety and anonymity of the users as well as to be protected against any kind of attacks that may be taken over the data in the channel.

8.1 Framework Features

8.1.1 Federation Establishment

Based on what was presented on the section 6, SAML specifications are the best choice to achieve a federated environment. Hence, SAML is the federated initiative that the GlobaliD framework addresses for establishing identity federation.

SAML provides several bindings to accomplish identity federation, creating multiple different ways for systems federate identity. Nonetheless, there are bindings more secure than others. Hence, developers must implement all the available SAML bindings for establishing identity federation because thereby the required security for establishing federated identity is asserted by which binding the service providers choose for exchanging messages with the GlobaliD identity providers. Nevertheless, the GlobaliD identity providers should have the artifact resolution binding end point by default, because this binding is the most secure binding that SAML defines.

The GlobaliD framework makes use of the SAML Web Browser SSO profile guidelines for establishing the communications between systems, where the user browser is used as an intermediary for the transference of SAML messages.

The SAML identifier chosen must be either the persistent or transient SAML identifier according to the context for the effect. When service providers do not request any information from the users but just an authentication context assertion about the user, a SAML transient identifier should be used instead.

A GlobaliD identity provider should supply the authentication context according to the last authentication method the user used when he/she authenticated at the GlobaliD federated identity provider, unless the service provider requires a specific one that was not used by the user on his/her last authentication at the GlobaliD IP. In such case the GlobaliD should re-authenticate the user according to the requested authentication context or issue an authentication failed response if necessary.

Pseudonyms must be used for the representation of users at the services providers they federate with. Users' anonymity is thereby assured (see Figure 15).

All SAML messages should be signed and encrypted by the issuer parties involved in the federation of the users' identity.

The framework also demands the implementation of the Single Logout mechanism in the identity providers based on this framework. This mechanism allows users to near-simultaneous log out from all the services they are using at a given time.

8.1.2 Associations

Each GlobaliD digital identity must be associated to several users self-intrinsic characteristics and to a few devices they use every day as a certification of the users' digital identity. The association between the users and their GlobaliD digital identity must be established by using (combination of two or more) of the following factors:

- The Citizen Card – gives several information of the holder (user) to the digital identity. Part or all the information retained in the users' citizen cards should be extracted to the user GlobaliD federated digital identity profile according to the user specific allowance. Such procedure assures that the user holding the GlobaliD really exists. Moreover, the digital signing functionality should be used as a certification of the users' GlobaliD digital identities.
- The finger print – gives a solid intrinsic unique identifier about the user. This is therefore another factor that should be used to assure the credibility of the user holding the GlobaliD digital identity.
- The email – it makes a relation between the identity and one of the users' means of communications online. This mean of communication is important since it represents the users body online as it was told in subsection 2.3 Representation of Digital Identity..
- The users' mobile – bounds the GlobaliD digital identities to a personal and real communication device.

The different kinds of association are established by the extraction of users' particular information from any of these elements and by using their functionalities to apply on the users' digital identity such as the authentication feature they have. These associations are excellent in the way they assure the federated services that the users really exist, they provided real information to the digital identity and so users are accountable for the actions they take. Moreover they are a great protection against identity theft.

Several multi-service providers as Google were several times convicted by a court of law to pay compensation for an offensive content to a person that was uploaded or made by someone by using the several services Google provides. By requesting GlobaliD identities associated to the users' citizen card, it would prevent this kind of lack of accountability a significant number of users tend to commit for they would be easily identified and consequently blamed.

8.1.3 Authentication

The GlobaliD framework demands the implementation of (one or more) the following users' publicly available authentication mechanisms in any GlobaliD Identity provider: password-based, email, citizen card, finger print, mobile [74]. The methods referred previously, with the exception of the password-based mechanism, can be used to implement strong authentication in a GlobaliD identity provider. The benefit of having so many authentication mechanisms is that when one is not available the other may replace it. For example, the users' finger prints may be used to authenticate users at the identity provider when their citizen card was either renewed or revoked by the state or even when they lost it. In such cases the association factor must be no longer valid and the holder must not be capable to use it to authenticate at the GlobaliD IP. Therefore, users should be provided a way to renew the association of the renewed authentication factor when this scenario arises. Users could use any of the other authentication means to access their digital identity in the meanwhile they do not have replaced the lost one. The same applies to the mobile, finger print and email factors when used. One important association factor besides the citizen card one is the users' mobile. It is relevant because such could be an authentication factor when users access websites through smart phones [74].

8.1.4 Profiling

GlobaliD demands the provisioning of a relatively simple interface by which users can organize their personal data and selectively share it with web applications and others (identity federation establishment). The users' particular information is managed via profiling it in a digital identity.

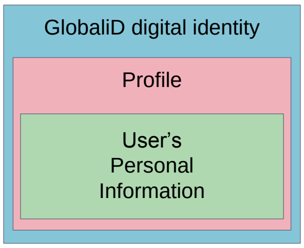
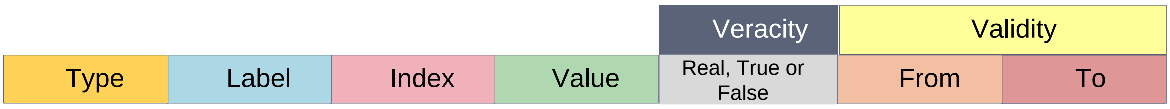


Figure 21 - Digital Identity Information.

Figure 21 shows that each GlobaliD digital identity only has/is one profile, and the profile has the personal information of the holder of the digital identity.

8.1.5 Information description

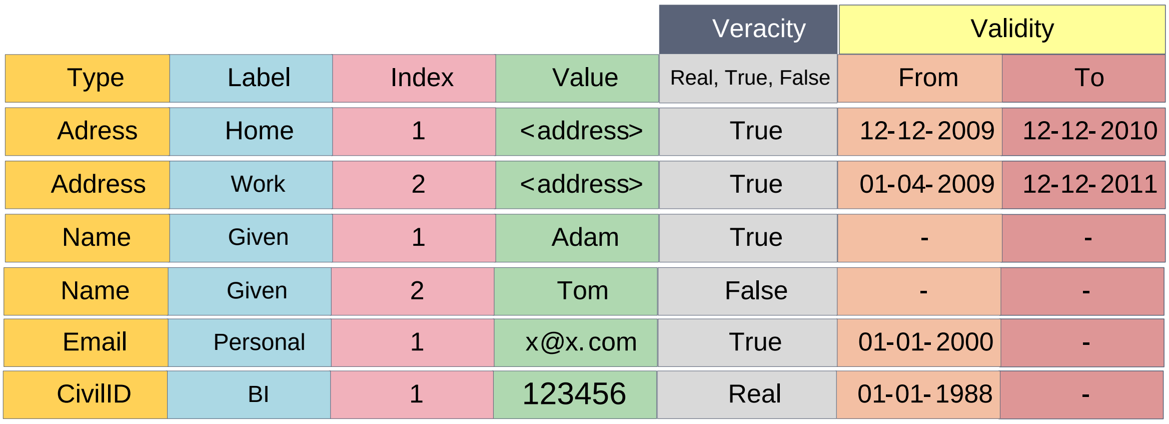
The information users provide to the profile may be of several types, such as Name, Date of Birth, Age, Address, etc… Usually users tend to have several pieces of the same type, for example: Home address, Work address. Thus, GlobaliD demands the following framework for describe the type of each piece of information:



				Veracity	Validity	
Type	Label	Index	Value	Real, True or False	From	To

Figure 22- Information description model.

Type is the kind of the information: Address, Phone, Name, date of birth. Label is the usual customized description of the type of information e.g., Home Address, Work Address. Index is a numerical order reference for the information of the same type and Label: Home Address 1 and Home Address 2. Value is the actual information. Validity is the information's time of effect. Veracity tells whether the information is real, true or false. The real veracity is only controlled by the GlobaliD systems, it is used to assert the veracity of the information obtained from the users publicly available means of authentication. Figure 23 provides an example of information stored in a GlobaliD identity.



				Veracity	Validity	
Type	Label	Index	Value	Real, True, False	From	To
Adress	Home	1	<address>	True	12-12-2009	12-12-2010
Address	Work	2	<address>	True	01-04-2009	12-12-2011
Name	Given	1	Adam	True	-	-
Name	Given	2	Tom	False	-	-
Email	Personal	1	x@x.com	True	01-01-2000	-
CivilID	BI	1	123456	Real	01-01-1988	-

Figure 23 - Information description example.

Figure 23 shows part of user's profile. It has several types of information which veracity value may be real, true or false. Label, Index and Value can be whatever users want. The user may use this information to create information cards.

GlobaliD makes available the following types for describing the personal information of users are: Address, CV, Date, Email, File, Finger Print, Gender, GPG, IM, Name, Nationality, Telephone, Photo, Profession, Status, Title, URL.

8.1.6 Information cards

The identity metasystem model described in the subsection 6.6 uses information cards to conduct the information users send to the web applications. For the metasystem, an information card is a representation of a digital identity, which is a group of pieces of information related to a given user. The information card concept (an analogy to the several cards we use in our real life) is a good idea for structuring the users' personal information. Therefore the GlobaliD framework demands the use of the information cards concept for the organization of the particular information users want to share with web applications and others. Thereby users will send their information to service providers or share it with others identities by gathering it in a group named information card. Users may name each information card they create. By doing so as a means to organize the users' information, users will be capable to selectively choose only the pieces of information they want to send to the services providers when federating with them or to share it with others users. Moreover it is an easy way to organize the information to send to the services. Information cards are a resemblance of personae by the GlobaliD framework.

Figure 24 shows the GlobaliD digital identity and the information cards hierarchy with examples of possible information cards that a GlobaliD may have.

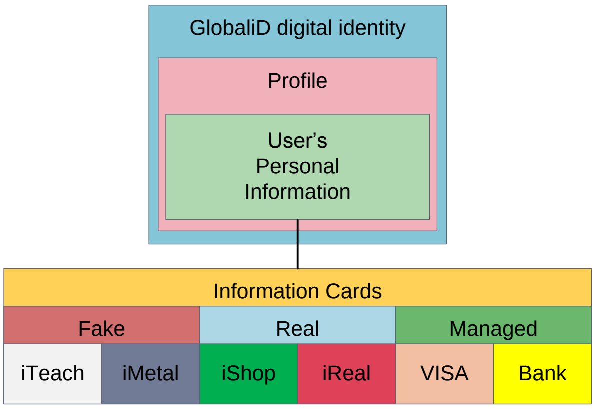


Figure 24 - GlobaliD digital identity and the information cards hierarchy.

In the Figure 24 the user has several kinds of information cards. The cards are divided in three base categories Fake, Real (SelfCards) and Managed.

Figure 25 provides an illustration of the information cards structure of the GlobaliD's.

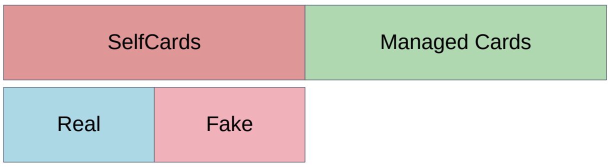


Figure 25 - Information cards framework.

In the figure it is illustrated the structure of the information cards by the GlobaliD. There are two types of information cards:

- Self cards – cards created by the user. Self cards may be either Real or Fake.
  - Real – Real cards only have real information. The real information is inherited from the user's citizen card.
  - Fake – Fake cards may have real, true and false information.
- Managed cards – cards asserted to the user by an outer entity in order to enable him/her to access or to use any of the services of the entity that asserted the managed card.

The GlobaliD IP may create an iReal information card automatically from the users' citizen card information of the user.

In the case of the GlobaliD being associated to the user's CC, because the information card they provided for federating with the web application is indirectly associated with it and with a pseudonym, users will be more reluctant at the moment of committing any action that may harm someone or fall in any kind of crime. In the case that such a thing would happen anyway, the user that committed the unacceptable action would be the one accountable for it and not the service provider. The use of pseudonyms for the representation of users in the services has the additional value of making users accountable and anonymous at the same time. This way the GlobaliD obtains accountability without giving up about either the users' privacy or anonymity.

Example

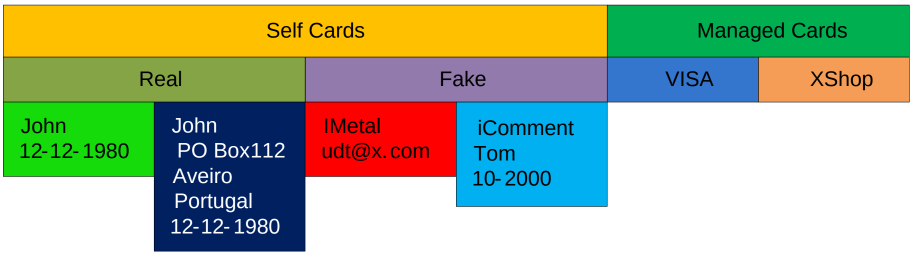


Figure 26 - John information cards.

Figure 26 shows a possible information cards scenario. In this case, the real user named John has two types of information cards, self cards, created by him (asserted by the Identity provider) or managed cards, asserted by the specific authority.

The self cards type is made by two subsections, Real and Fake. The real subsection holds the information cards that only have information inherited from the John citizen card or any other means of associations referred previously. Therefore the real section holds information cards with only true claims, certified by an external authority. Real self information cards may be provided to an e-business, for example an amazon.com alike, when purchasing products in order to smoothly supply the service with the buyer's addressee. The fake section holds information cards that have either real, true or fake information. Information cards of this kind can be used to associate with web applications of social engagement as blogs, forums, and social networks in general. Users will be protecting their identity when doing so. The Managed cards section holds information cards asserted by other entities such as the one shown in the figure, VISA, in order to enable the user to pay shops online.

8.1.7 Information classification

GlobaliD users are given the possibility to manually provide other information to their profiles beside the one extracted from the associations factors mentioned previously. The information given by them may be asserted by either as true or false. In order to tell apart which true information of the profile is provided by any of the users' association factors and which is manually supplied by the user, a numerical ranking mechanism is created to assert about it. The ranking mechanism has three levels from 0, 50, 100, where 0 means that the information is false, 50 means that the information is asserted as true by the user and 100 means that the information is true and consequently it was obtained by any of the users’ association factors. Thereby, the services will be able to discern which true information from the users was provided by their citizen cards and which one was provided manually by them. GlobaliD asserts the veracity ranking of an information card through the following formula:

$$VR(I.C.) = \frac{V_{Real} \times \sum_{i}^{n} i_{Real} + V_{True} \times \sum_{i}^{n} i_{True})}{n \times V_{Real}} \times 100$$

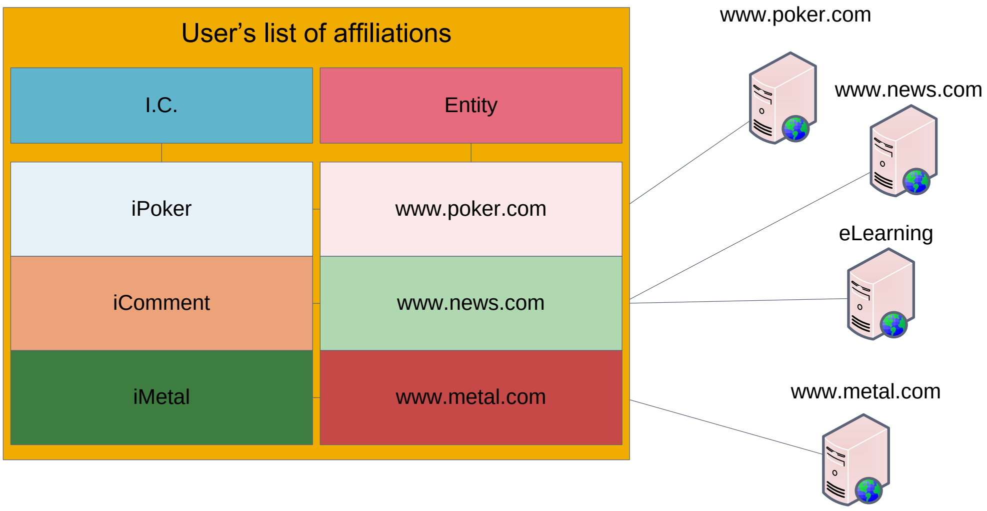
Figure 27- The information card veracity formula.

In the formula the VR is the veracity ranking of the information card. The veracity ranking output range is from 0 to 100 percent. VReal and VTrue are the values of 100 and 50 of veracity of the respective piece of information; ‘i’ is the respective piece of information.

The services providers may create their own kind of rankings over the information cards users assert to them. In the case of a news commenting website, services may provide users with a mechanism that allows them to evaluate each other’s comments, such as the very used thumb up thumb down one. Thereby the website could provide a list users classification (points).

8.1.8 Scrutiny and information displacement monitoring

GlobaliD users will share information with others, visit websites and register in them. Therefore the framework demands that all user visits to the websites, and consequently registration (affiliations) should be logged. The same applies when users share their information to any entity via any mean of communication. Thereby GlobaliD users may monitor the scrutiny and displacement of their personal information and consequently, the web sites they visit and web applications they provided information cards to, by keeping these events logged. Several kinds of lists for different purposes may be created. For example a list of affiliations will save information about which information cards were sent to web applications. The following figure illustrates it.



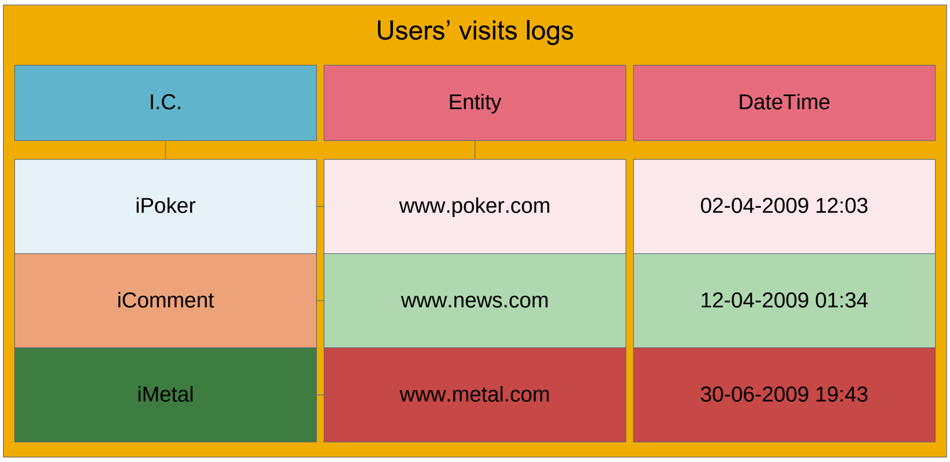
User's list of affiliations

I.C.	Entity
iPoker	www.poker.com
iComment	www.news.com
iMetal	www.metal.com

Figure 28 - User's list of affiliations.

This mechanism allows users to be always aware of the whereabouts of their personal information. When they consult the affiliation list they know to which website each information card was sent to.

Another list could keep the users' visits to websites as show in the Figure 29.



Users' visits logs

I.C.	Entity	DateTime
iPoker	www.poker.com	02-04-2009 12:03
iComment	www.news.com	12-04-2009 01:34
iMetal	www.metal.com	30-06-2009 19:43

Figure 29 - Logs of the visits users made to the specified website.

8.1.9 Information Privacy Regards

Information cards are groups of pieces of information. They are used to send to the web applications in order to established identity federation with it. Service providers should only save it in their database after the approval of the owner of the information (the user). When users do not allow services to keep their information, services should only save the pseudonym association for futures access to the service. In the case service will need personal information from the user in the future they must request the users to allow them to retrieve it from their IP via an information card assertion. All the information cards sent to the service providers must be encrypted.

8.1.10 Communication Channels

In order to provide the best security when exchanging the information among the identity provider and the service provider, the GlobaliD demands the implementation of the HTTP protocol over TLS for establishing secure communications with others systems. By doing so, the channels used to exchange users' identity information are protected against eavesdroppers, man-in-the-middle attacks, and identity theft. Consequently, it is protecting the users' privacy and anonymity. Moreover, mutual authentication between the systems is implemented, users authenticate to the services providers, but the service providers also authenticate themselves to users (having the GlobaliD IP as proxy).

8.1.11 User centric address book

GlobaliD users have their personal information stored in a profile. It would be interesting if they could share it with other GlobaliD users, as well. The GlobaliD federated digital identity also enables information sharing between users that have a GlobaliD identity even if the GlobaliD digital identities do not belong to the same GlobaliD Identity Provider. The information is shared via information cards. This way the users build a user centric address book, always up to date.

8.2 GlobaliD General Overview

Figure 30 makes an overall overview of the GlobaliD Framework from the user's point of view.

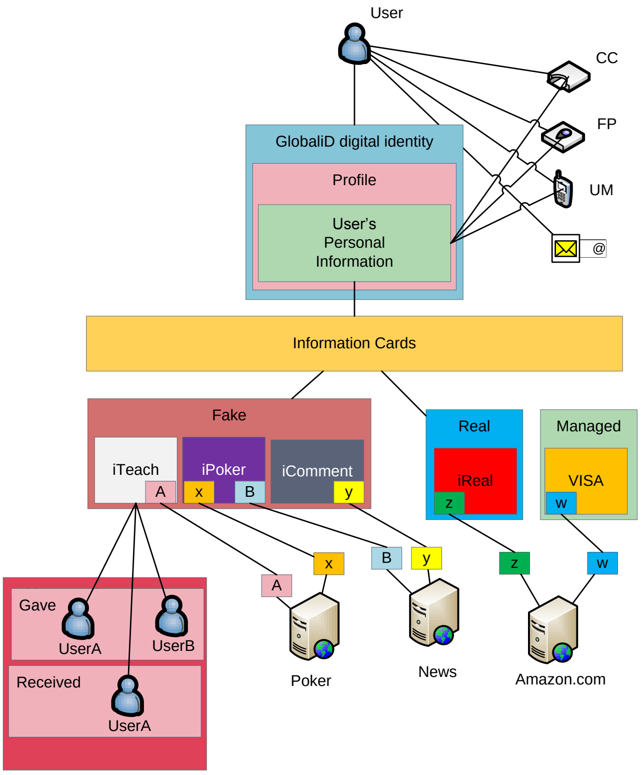


Figure 30 - GlobaliD Framework Overview.

The figure illustrates the GlobaliD Framework from a user point of view. The user has a GlobaliD digital identity. His GlobaliD Digital identity is a profile. The digital identity is associated to his citizen card (CC), finger print (FP), mobile (UM) and email (@). These associations provide particular information about the user to his GlobaliD digital identity profile. The user's particular information is used to create information cards to represent him differently in the web services. The user provides information cards to the web services in order to establish identity federation with them. The association between the user's information cards and the several services he is federated with is established by the pseudonyms. A per web application pseudonym is used as a representation of the information card at each the service provider. One Information card may be used in more than one service provider. The information card maybe exchanged with other users.

8.3 GlobaliD Federated Identity Environment

Figure 31 gives an overview of a GlobaliD Identity Provider in a GlobaliD identity federated environment. MiD and XiD are both GlobaliD Federated Digital Identity Providers.

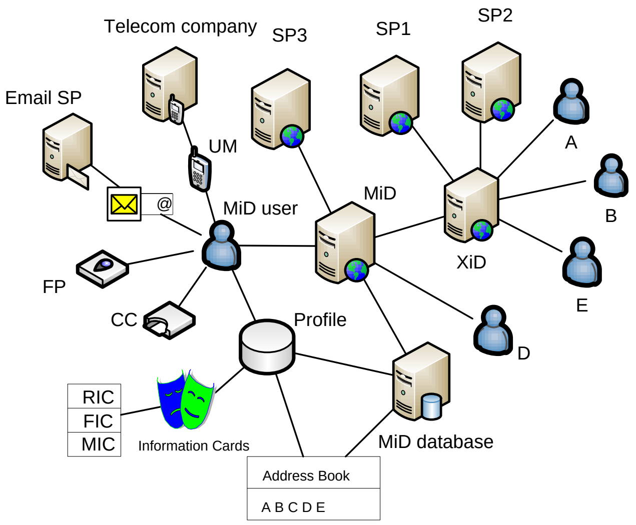


Figure 31 - GlobaliD Identity Provider framework.

The figure represents a set of user profiles in the MiD and XiD identity providers (IPs). The profiles are MiD user, A, B, C, D and E. All profiles are GlobaliD digital identities. The trustworthiness and security of the MiD user identity is taken higher by associating it with some characteristics and technologies the users owns, such as his citizen card, digital signature, finger print, email and mobile. The different kinds of associations are made by the extraction of users' particular information from any of these elements or by using some of their functionalities to apply on the users' digital identity. The citizen card gives several information of the holder (user) to the digital identity. This information is certified by the user national state. The digital signing functionality enables users to certify their digital identities by signing it. The finger print gives a solid identifier about the user. The email makes a relation between the identity and one of the users' means of communications online. The users' mobiles bound their digital identities to a personal and real communication device. By making these kind of associations (@, CC, FP, UM) the MiD IP is assuring the federated services that the user really exist and is accountable for the action that he takes. Moreover strong authentication at the MiD IP is implemented by using the authentication capabilities of the user's association devices.

Regarding the MiD user, part of the information of his profile was extracted from the devices he associated himself with, and kept marked as real information of the profile. The user can add more information to his profile and mark it either as false or true. The user can use his digital identity to federate himself with web applications (SP1, SP2, and SP3), using SAML assertions rendered by the MiD IP. The MiD user uses information cards to selectively send his personal information to the web applications he wants to federate with. In this case his GlobaliD IP provides him a pseudonym in order to establish the association between him and the web application he wants to use. According to the services demands, the user may use information cards of real (RIC) or fake information (FIC) to supply to the services he uses, and other cards (MIC) may be asserted by other authorities, such as banks and shops, to authorize the user to access their services, or to pay shops online, as a information card asserted by the user's bank would be nice to accomplish it if it would be possible. The fake information cards (FIC) can be used in websites that do not request real information from the users, such as the websites of news, in order to protect his privacy when commenting any news. The pseudonym association established when the user uses any kind of information card to federate with the service provider forces him to behave, since he may be easy accountable for the actions taken within the public areas of the service. By using the classification of true, fake and managed information for the information of the users' information cards, MiD creates its own ranking of trust over the users' personal information, and consequently over the digital identity itself, according to the GlobaliD framework demands. Moreover, the use of information cards allows the websites to create a reputation ranking over it.

The user may use information cards not only to selectively send information to web applications but also to selectively share it with other users or entities. Thereby, the user builds a user centric address book always up to date. Any GlobaliD IP represented implements the GlobaliD framework asserted security guidelines in order to protect the channels of communications against several kind of attacks.

8.4 Major benefits

The GlobaliD framework provides:

- Smooth digital identity management – by providing users with an interface for managing their personal information according to each piece of information.
- Seamless access to web applications – by providing single sign on capabilities to their digital identities.
- Near-simultaneously log out of all the applications users were authenticated at – by providing a single log out mechanism.

- Secure channels of communications – by implementing the use of HTTP over TLS. This way the channels used for exchanging the users' personal information are encrypted, avoiding the unauthorized disclosure of any user's data.
- A structure to organize the information sent to web applications – by grouping the users' personal information in information cards.
- Scrutiny and displacement monitoring – by creating a list of affiliations with websites and personal information receivers.
- Control over the disclosure and dissemination of each piece of the users' personal information – by delegating to users which information about them they desire to provide to the web applications.
- Integrity and validity assured in information exchanging – by asserting that the SAML messages should be signed and encrypted by the parties involved in the federation of the users identity.
- Information classification (real, true, false, and managed) – by providing a mechanism to assert about the veracity of the information they provide to their GlobaliD digital identity.
- Strong and mutual authentication mechanisms – by using the association factors to implement strong authentication at the GlobaliD identity provider and by demanding the implementation of HTTP over TLS on the channels of communications, which permits that system authenticate to users as well.
- Digital identity certification and classification – by the use of the users' publicly available authentication mechanisms and self-characteristics to associate to their GlobaliD digital identity and by the provisioning of an information card veracity ranking over the information it holds.
- Protection of privacy and anonymity – by demanding the use of pseudonyms to represent the users in the identity federation association.
- Simultaneous anonymity and accountability – by the associations of the users' publicly available means of authentication to their GlobaliDs.
- User centric address book always up to date – by allowing that users share their personal information with other users that own a GlobaliD digital federated identity as well.

8.5 Major flaws

- GlobaliD maintains all the drawbacks of being a federated model.
- Users may fear the association of their digital identity with their citizen cards (or to any of the mentioned associations). This is high influenced by the user cultural context.
- Personal information concentration may also be of concern (the big brother effect).
- After users share their personal information to a service provider or to other users they cannot control its replication or provisioning to third parties.

8.6 State of Art Comparison

The GlobaliD Federated Digital Identity framework goals are the protection of the users' privacy, anonymity, the confirmation of its veracity as well as to make them reliable and accountable about the actions they take on the virtual world. The following table compares several initiatives discussed in this paper with the GlobaliD Federated Digital Identity features.

Table 4 - Federated Initiatives Comparison.

	OpeniD	**W.L. ID**	**I.C Metasystem**	**WS-***	**SAML**	**GlobaliD**
Unique identifiers	Yes	No	No	Yes	No	No
Pseudonymity	Not demanded	Yes	Yes	Not demanded	Yes	Yes
Information categorization **(R/F/Managed)**	No	No	No	No	No	Yes
Personae	Not demanded	No	No	Not demanded	No	Yes Via I.C.
Information validation	No	No	No	No	No	Yes
Strong Authentication Mechanisms	Not demanded	No	Yes (P.I.N.)	No	Yes	Yes
Identity ranking	No	No	No	No	No	Yes

By the analysis of the Table 4 it may be asserted that the GlobaliD framework has features that strengthen the users, services and the whole federation, taking the federation identity management to a higher level. Instead of using unique identifiers for the representation of users' digital identities, the GlobaliD uses the SAML pseudonyms. It avoids the discerning of users by the correlation of identity identifiers and protects the privacy and anonymity of users as well. GlobaliD categorizes the information of users in terms of false, true and real data, demanding the use of information cards to represent it. This way, users can organize the information they send to the web applications they want to use and also monitor its whereabouts. The information classification permits the creating of several rankings (veracity, behavioral) over the users' information cards by the GlobaliD identity providers as well as the services providers may create their own rankings over the information cards users provide to them. By using the users' citizen cards or any other users' publicly available means of authentication, GlobaliD validates their information, their identity and implements strong and mutual authentication mechanisms at the identity providers, as well. In the absence of any of the users publicly available means of authentication they may use the traditional password-based authentication to access their GlobaliD digital identity, nevertheless. The association of the users' digital identities with their several self-intrinsic characteristics and devices they use every day makes users credible, reliable and accountable about the actions they take in the services they use, preventing their misbehavior in the first place. In any case they are always accountable for any action taken by them.

8.7 Conclusion

The GlobaliD framework smoothes and improves the users' digital identity management online by providing:

- Seamless access to the web applications.
- Control over the disclosure and dissemination of the personal information.
- Scrutiny and displacement monitor over the information shared.
- Classification of the information shared as real, true, false and managed.

It also manages to strengthen the users' privacy and anonymity in their association with the service providers by:

- Implementing pseudonymity for the representation of users.

Consequently, the use of pseudonyms for the representation of users makes them accountable at the services they use without giving up about the anonymity of the users.

GlobaliD also demands the implementation of secure protocols that assure the protection of the channels of communication from several kinds of harmful attacks. By using publicly available strong authentication means it assures the veracity of the digital identity information, its authorized access and prevents identity theft in the first place. It makes users accountable about the actions they take on the web applications without giving up on their anonymity as well. Users may have a user-centric address book always up to date. The SPs are set free from the identity management of their users, decreasing their maintenance costs and improving their services in the first place.

Even though that the GlobaliD framework asserts security principles to the systems, the security of the digital identities are also a task that must be made by their holders, in order to avoid ill intend services, which may get their information for harmful purposes. Users also must assure whether the service provider is a reliable one by checking the browser identity information.

As a general analysis, the GlobaliD framework fulfills the initial goal of taking a step further in the digital identity management world by making it more versatile, responsible, trustworthy, integrity and privacy safe, anonym and thus secure.

9 The GlobaliD Identity Provider

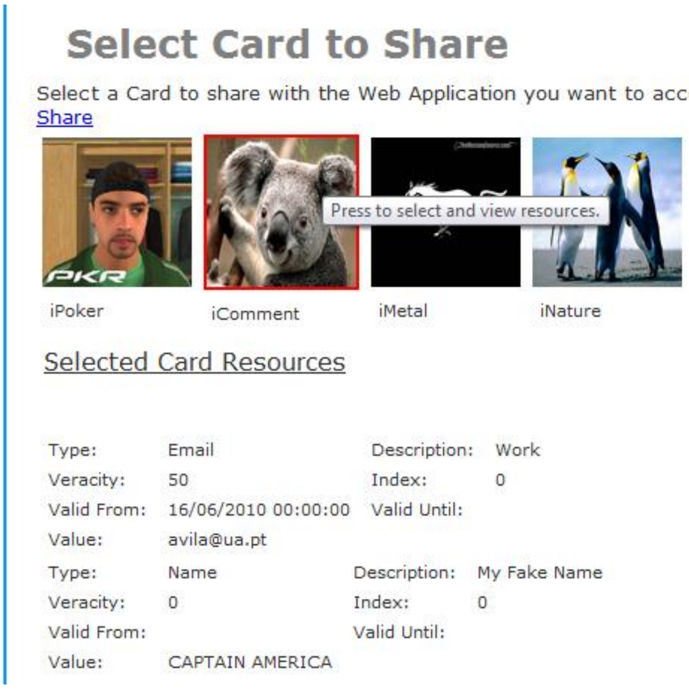
In order to demonstrate a practical example of the GlobaliD, a GlobaliD federated identity provider prototype named GlobaliD was developed. This prototype, as the prototyping software engineering model dictates, provides limited functionalities of the GlobaliD framework. It also helps on finding more use cases to implement in a future solution for production. The aim for the development of this prototype is to prove that the management of the users' digital identities may be improved in comparison to other federated identity frameworks by using only part of the GlobaliD framework identity specifications, and thereby the digital identity management is fairly enhanced.

9.1 Federation

The GlobaliD IP follows the SAML specifications to establish federation between the users and the web applications as defined by the GlobaliD framework specifications. However only some features were implemented:

- Only the SAML HTTP POST binding was implemented.
- The single log out mechanism was not implemented.
- Only the transient identifier is used to refer to the format of the users' identifiers. Thereby users will have to indicate which information card they want to share at each federation engagement (each website visit).
- SAML pseudonyms are used for the representation of the users.
- SAML Messages are signed and encrypted.
- Only the association of the GlobaliD with the users' citizen card is integrated in the prototype.

Figure 32 shows the page to which the user (testuser) is redirected when she wants to federated with a web application.



Select Card to Share

Select a Card to share with the Web Application you want to access to.
Share

Press to select and view resources.

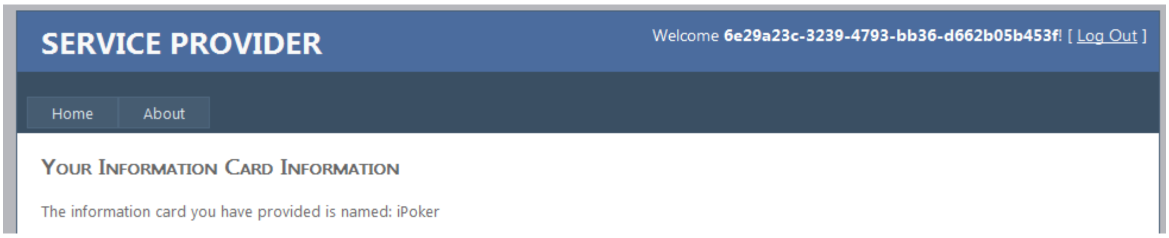
Selected Card Resources

Type:	Email	Description:	Work
Veracity:	50	Index:	0
Valid From:	16/06/2010 00:00:00	Valid Until:	
Value:	avila@ua.pt		
Type:	Name	Description:	My Fake Name
Veracity:	0	Index:	0
Valid From:		Valid Until:	
Value:	CAPTAIN AMERICA		

Figure 32 - Federation Page.

On this page the user can select which information card she wants to share with the web application (service provider) she is about to federate with.

Figure 33 shows the information card shared with the SERVICE PROVIDER web application. On the right top corner, right after the Welcome word it is displayed the pseudonym asserted to the information card provided by the testuser at the time of identity federation. Since the prototype is only asserting SAML transient identifiers the user test has a new pseudonym each time she accesses the web application, even if she already have accessed it previously, and no information card is stored by the SERVICE PROVIDER web application. If a SAML persistent identifier would be used instead, the web application would store the information card in their database, and its pseudonym would be used to refer the user in future visits (per-web application pseudonym). The user provides an information card each time she wants to access the web application by using her GlobaliD within the federated context of this prototype. The information cards sent to the web applications will be encrypted.



SERVICE PROVIDER

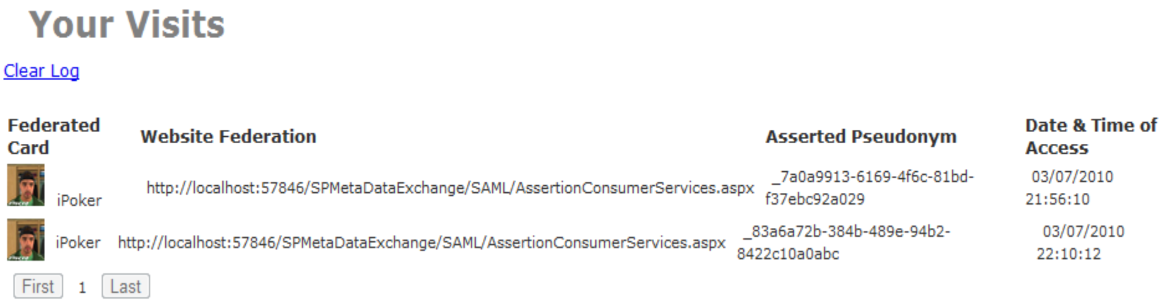
Welcome **6e29a23c-3239-4793-bb36-d662b05b453f**! [Log Out]

YOUR INFORMATION CARD INFORMATION

The information card you have provided is named: iPoker

Figure 33 - User Information at SERVICE PROVIDER.

In order to enable users to know which web applications they have federated with until their present time, the GlobaliD IP prototype provides a logging mechanism for the purpose. Figure 34 shows the testuser past identity federation. In each identity federation it is recorded which information card was provided to the federated web application, which web application the information card was provided to, which pseudonym was asserted to the information card and the date and time the identity federation took place.



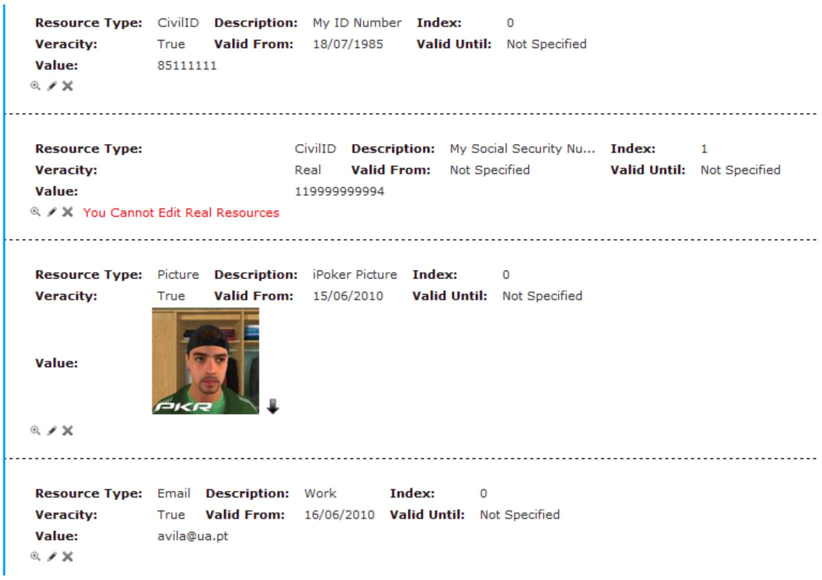
Your Visits

Federated Card	Website Federation	Asserted Pseudonym	Date & Time of Access
iPoker	http://localhost:57846/SPMetaDataExchange/SAML/AssertionConsumerServices.aspx	_7a0a9913-6169-4f6c-81bd-f37ebc92a029	03/07/2010 21:56:10
iPoker	http://localhost:57846/SPMetaDataExchange/SAML/AssertionConsumerServices.aspx	_83a6a72b-384b-489e-94b2-8422c10a0abc	03/07/2010 22:10:12

Figure 34 - User past identity federations.

9.2 Information management

An interface was developed for the management of the users' information according to the GlobaliD guidelines for personal information description. Figure 35 shows a print screen of the application information management profile. In the figure it can be seen a few resources from an existing user named testuser. Each resource shown is formed according to the GlobaliD specifications.



Resource Type: CivilID **Description:** My ID Number **Index:** 0
Veracity: True **Valid From:** 18/07/1985 **Valid Until:** Not Specified
Value: 85111111

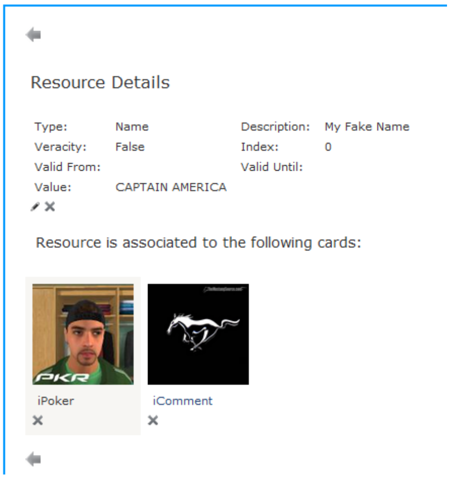
Resource Type: CivilID **Description:** My Social Security Nu... **Index:** 1
Veracity: Real **Valid From:** Not Specified **Valid Until:** Not Specified
Value: 119999999994
You Cannot Edit Real Resources

Resource Type: Picture **Description:** iPoker Picture **Index:** 0
Veracity: True **Valid From:** 15/06/2010 **Valid Until:** Not Specified
Value:

Resource Type: Email **Description:** Work **Index:** 0
Veracity: True **Valid From:** 16/06/2010 **Valid Until:** Not Specified
Value: avila@ua.pt

Figure 35 - Sample of some testuser's resources.

Each resource can be deleted by clicking the cross icon or edited by clicking on the pencil icon. If the resource is of veracity Real, it cannot be edited because this resource was extracted by some public available authentication mechanism of the user and therefore cannot be modified. Nevertheless it can be deleted because the user may wish not to have this kind of information stored in his profile. When the magnifying glass icon is clicked the user is redirected to a page where all the resource's details are listed and which information cards is that resource associated to. Figure 36 shows which information cards a testuser's resource of type Name and value CAPTAIN AMERICA is associated to.



Resource Details

Type: Name Description: My Fake Name
Veracity: False Index: 0
Valid From: Valid Until:
Value: CAPTAIN AMERICA

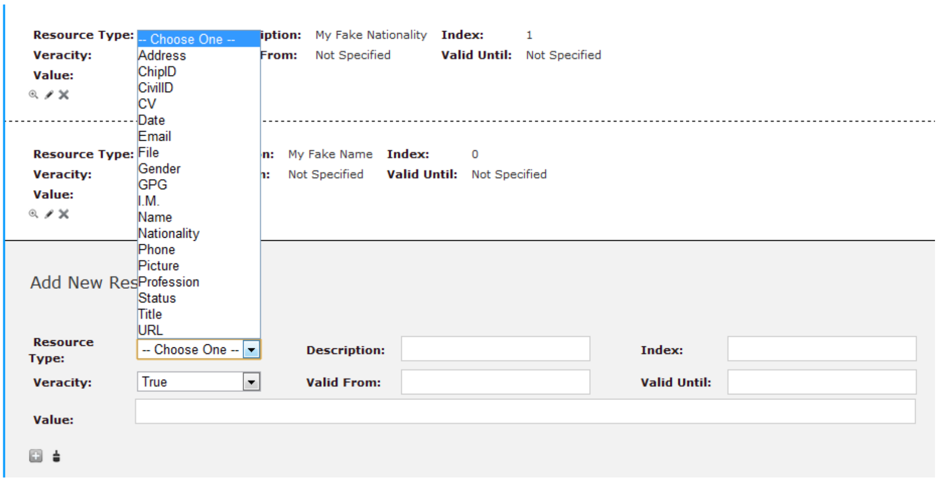
Resource is associated to the following cards:

iPoker
iComment

Figure 36 - Resource associated to the cards shown.

The resource shown is associated to the testuser information cards iPoker and iComment.

User may add new resources to their profile and may assert its veracity as either true or false by using the form for it. Figure 37 shows the form addressed to add new resources to a profile.



Resource Type: -- Choose One -- iption: My Fake Nationality Index: 1
Veracity: From: Not Specified Valid Until: Not Specified
Value:

Resource Type: n: My Fake Name Index: 0
Veracity: n: Not Specified Valid Until: Not Specified
Value:

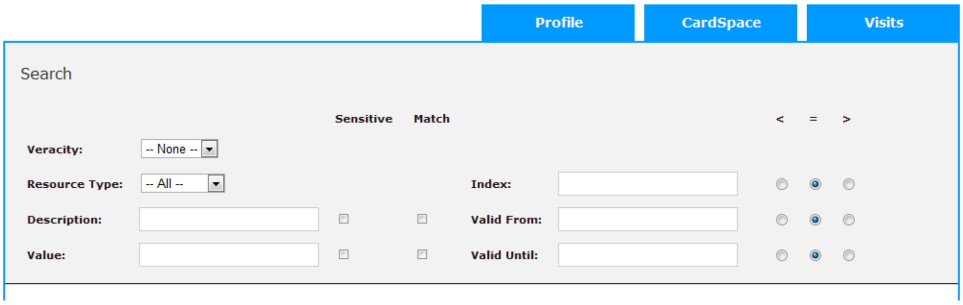
Add New Res

Resource Type: -- Choose One -- Description: Index:
Veracity: True Valid From: Valid Until:
Value:

Figure 37 - Form for add new resource to the profile.

Users can add new resources of any of the types indicated by the Resource Type drop down list. The form will adapt to the new type of resource users select e.g., if the user wants to upload a new picture the form will display an upload control instead of the textbox shown in the value row.

User may also look for a particular resource or group of resources with specific features by using the search form illustrated by Figure 38 .



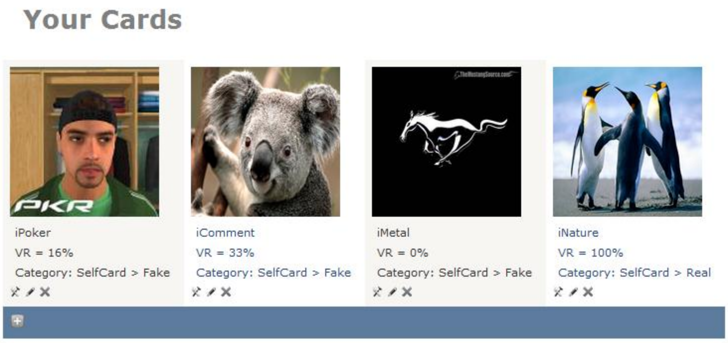
Search

		Sensitive	Match			<	=	>
Veracity:	-- None --							
Resource Type:	-- All --			Index:		○	●	○
Description:		□	□	Valid From:		○	●	○
Value:		□	□	Valid Until:		○	●	○

Figure 38 - Search form.

9.3 Information Cards

It will be possible for users to create information cards from the resources they had provided to the profile. Managed information cards specified by the GlobaliD identity framework are not supported. Figure 39 shows the testuser Cardspace.

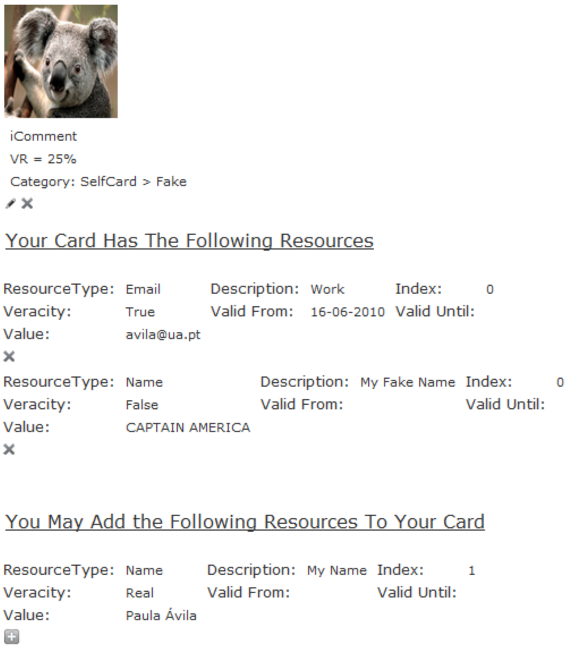


Your Cards

iPoker	iComment	iMetal	iNature
VR = 16%	VR = 33%	VR = 0%	VR = 100%
Category: SelfCard > Fake	Category: SelfCard > Fake	Category: SelfCard > Fake	Category: SelfCard > Real

Figure 39 - testuser Cardspace.

The user has four cards in her Cardspace. Each CardSpace has a picture, an alias, a veracity ranking indicated by VR and is integrated within a kind of category. In the case of the iNature information card it has a VR of 100% meaning that information card has only real resources and therefore its category is the Real one. The following picture shows the iComment resources and how resources are added and deleted from any card.



iComment
VR = 25%
Category: SelfCard > Fake

Your Card Has The Following Resources

ResourceType:	Email	Description:	Work	Index:	0
Veracity:	True	Valid From:	16-06-2010	Valid Until:	
Value:	avila@ua.pt				

ResourceType:	Name	Description:	My Fake Name	Index:	0
Veracity:	False	Valid From:		Valid Until:	
Value:	CAPTAIN AMERICA				

You May Add the Following Resources To Your Card

ResourceType:	Name	Description:	My Name	Index:	1
Veracity:	Real	Valid From:		Valid Until:	
Value:	Paula Ávila				

Figure 40 - Add or delete resources to an information card.

The respective headings indicate which resources the information card displayed has and which can be associated to it (all the resources user have on her profile). User can add and delete resources to and of the information card by just clicking on the plus or cross icons, respectively.

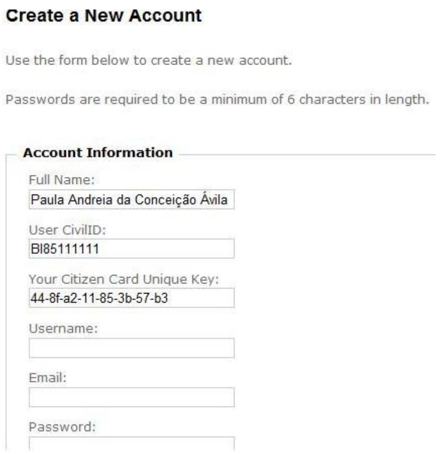
9.4 Information card ranking

Each information card is provided its veracity ranking based on the veracity of each piece of information it holds as specified by the GlobaliD framework guidelines. Figure 39 shows the veracity ranking of each card in the testuser's Cardspace. When a card has 100% of veracity ranking it means that all the resources contained in that information card were extracted from any of the public available means of authentication referred by the GlobaliD specifications. In the case of this GlobaliD prototype they could only be extracted from the testuser's citizen card. When a card has 0% of veracity it means it only has false information. Values between 0% and 100% indicate that it contains true and false information (user's resources), therefore it is a fake information card.

9.5 Association

Regarding the strong authentication factors defined by the GlobaliD framework, the GlobaliD prototype implements the use of the users' citizen cards to authenticate at the identity provider, but only the Portuguese Electronic Citizen Identity Cards are accepted.

No other means of associations specified by the framework are used. Therefore the real information shown in the user's profile could only be extracted from her citizen card. Figure 41 shows the GlobaliD prototype's registration form when the user chooses to register at the GlobaliD by using her citizen card. The user's full name, civil id and her citizen card serial number were extracted from the authentication certificate contained in her citizen card. These pieces of information are stored in her profile database as a certification of her GlobaliD and for automatic login when she is accessing the GlobaliD digital identity with her citizen card.



Create a New Account

Use the form below to create a new account.

Passwords are required to be a minimum of 6 characters in length.

Account Information

Full Name:
Paula Andreia da Conceição Ávila

User CivilID:
BI85111111

Your Citizen Card Unique Key:
44-8f-a2-11-85-3b-57-b3

Username:

Email:

Password:

Figure 41- Registration form when registering with the Portuguese Citizen Card.

9.6 Authentication

The users citizen card will be associated to the users GlobaliD Digital identity as a means for identity certification. It also will be used to automatically authenticate the users at the GlobaliD identity provider. No illustration his provided since the prototype redirects the user to her profile after a successful authentication took place.

9.7 Environment of the GlobaliD IP

Figure 42 illustrates the GlobaliD identity provider in a GlobaliD Federated Identity environment.

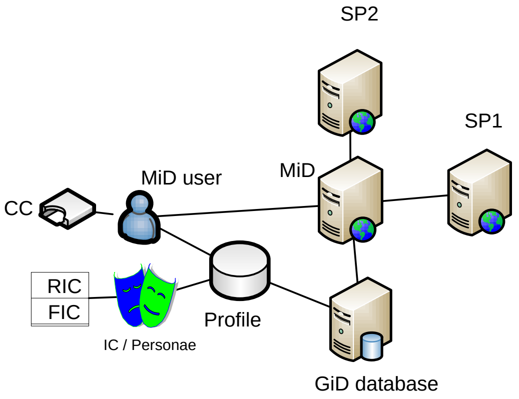


Figure 42 - GlobaliD IP Federated Identity Environment.

In the GlobaliD federated environment the user digital identity is only associated to her citizen card. The user can create only real and fake information cards. The managed cards specified by the GlobaliD framework are not possible to create. In order to federate with service providers the users supplies them with information cards as defined by the GlobaliD framework.

9.8 Database diagram

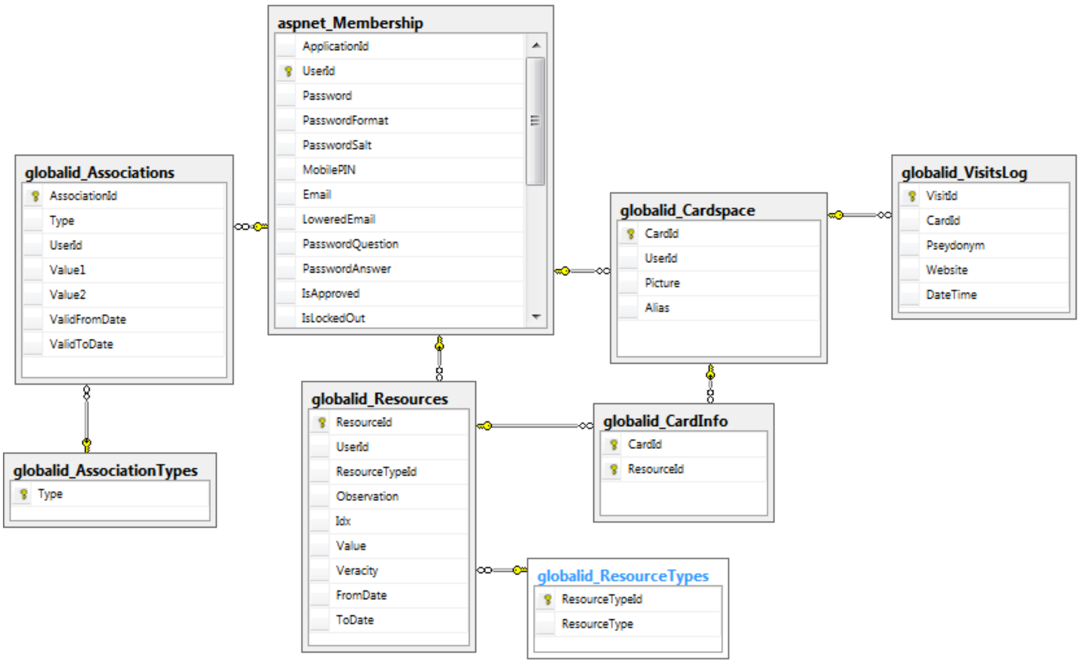


Figure 43 - GlobaliD Identity Provider Database diagram.

The diagram illustrated by Figure 43 provides an inner sight of the GlobaliD IP prototype database. The table globalid_Associations serves for keeping the necessary data of the users' publicly available means of authentication when they associate it with their digital identities for future automatic logins. The data collected is also needed for the certification of the users' digital identities. The globalid_AssociationTypes Table defines the type of the available users' associations (Citizen Card, Email, etc.) referred by the GlobaliD framework. The Table globalid_Resources holds the several types of information the users provide to their digital identities. The type of each resource is defined by the globalid_ResourceTypes table. Globalid_CardSpace represents the information cards users create and the globalid_CardInfo table holds the resources (users' personal information) of each information card. The table globalid_VisitsLog, is used to log users identity federation engagements. The aspnet_Membership table is the table used by the application to hold particular user account information.

9.9 Conclusion

The GlobaliD identity provider offers a very good illustration of the implementation of a GlobaliD federated identity context by using only part of the set of specifications defined by the GlobaliD framework. By implementing only the association of the users’ citizen cards with their digital identities, a fair amount of trust and accountability is obtained without giving up about the users' anonymity. The use of fake and real information cards grants the necessary certification about the users’ when they federate with the services. The interface provided by the prototype application offers users an easy way to manage their particular information. A simple and versatile interface for creating and deleting information cards is also offered. The application also makes available a logging mechanism for users knowing which entities they have shared their information with, and consequently the federations to which they belong. Registration at the prototype by using the Portuguese citizen card is implemented and therefore it is possible using it to login at the IP, which prevents identity theft.

10 Conclusions and Future Work

This Master’s thesis was addressed to the digital identity management discipline. Its goal was to take a step further in the digital identity management field by making it more versatile, responsible, trustworthy, integral and privacy safe, anonym and thus more secure.

In order to accomplish the goal, a profound study of the fundaments of the identity and digital identity management subjects was made. Their features and particularities were introduced, described and analyzed. By analyzing two models of digital identity management one had a clear comprehension of the procedures involved in managing the users’ personal information and their access to the service providers. Thereby it was concluded that the federated identity model is the most suitable model for digital identity management and users' access to the service providers since it efficiently delineates the respective concerns of each entity involved in the identity and access management process and grants high versatility, reliability and security to the users and the involved federated members. By using a federated identity users are capable of monitoring the provisioning and scrutiny of their information and logging all their activities online while having seamless access to web applications. The scattering of their information across the web is controlled and their anonymity and privacy remains unaffected as well. The main inconvenience of the federated identity management that is worth mentioning is that once an attacker gains admission to a federated digital identity, he/she/it is capable of accessing all the services the identity is federated with and gathering all the identity information. Nevertheless, such inconvenience can be easily avoided by implementing strong authentication mechanisms and encrypted channels of communications. Yet users are the first agents regarding the protection of their privacy and discernment. According to which was just pointed out, when projecting digital identity management systems they should be based on federated identity initiatives since it entails major advantages and not so significant hassles in comparison to non-federated identity solutions.

A federated identity framework named GlobaliD was presented based on the study made of the digital identity management discipline. The several parts of the framework were introduced, described, explained and analyzed. Based on the analysis of the GlobaliD and on its comparison to other federated identity initiatives available online, it may be concluded that the GlobaliD entails significant advantages considering the users' information management, privacy and security concerns, making it a relevant and valuable framework for digital identity management. The factors of association (publicly available means of authentication) defined to bind the users to their digital identities are strongly beneficial since they provide an excellent level of trustworthiness and security to their identities and make them consecutively more responsible as well as more accountable for the actions taken on the services they federate with. Nevertheless, this may make them reluctant to adopting a GlobaliD. Furthermore, the fact that the GlobaliD framework gathers relevant and accurate information (big brother effect) about the users and their activities may lead them not to adhere to the GlobaliD due to privacy and accountability concerns or lack of confidence in the GlobaliD identity provider.

In order to demonstrate the GlobaliD framework functionality, the Global Identity Provider prototype was developed. The development of this identity provider prototype proved that it is possible to supply the users and the federated services with highly trustworthy and reliable digital identities while keeping the users’ privacy and anonymity safe and still guaranteeing the seamless access to the web applications by using only specific parts of the GlobaliD architecture.

Even though the GlobaliD federated identity framework achieves a high level of trust and security and provides an enhanced solution for digital identity management, future work for improvements and adjustments of the framework will always be necessary according to the future requirements of digital identity management. Future work may delineate an association of GlobaliD to the users’ VISA enabling them a smooth online purchase (managed card assertion). Thereby, users would be capable of recharging their GlobaliD with a particular amount of money to directly pay products online by using their GlobaliD. Future requirements may be simply integrated since the framework is flexible and easily scalable. In fact, the GlobaliD framework may be used together with other federated initiatives with no problems. An integration of the GlobaliD with the metasystem information cards is possible through a simple adaptation, such as using the SAML SOAP Binding messages body to carry the information cards.

It can be said that the GlobaliD federated identity framework is a very straightforward solution for digital identity management because it is versatile and secure, whilst assures users’ responsibility, trustworthiness and the safety of their privacy and anonymity as well as the integrity of their data when exchanging it with other entities online. Therefore, the aim of bringing up an enhanced solution for digital identity management which could overcome the drawbacks of the existing models and initiatives for managing the users' identity and their access to the services providers in a reliable, trustful and subtle manner was achieved by the GlobaliD.

Finally, with the creation of the GlobaliD the goal of taking a step further in the digital identity management field by making it more versatile, responsible, trustworthy, integral and privacy safe, anonym and thus more secure was accomplished.

Bibliography

[1] K. Chen, "Protecting Personal Infomation Online: A Survey of User Privacy Concerns and Control Techniques," *The Journal of Computer Information Systems* 2004.

[2] F. a. foll and J. Baragry, "Next Generation of Digital Identity," *Telektronikk,* vol. 103, p. 4, 2007.

[3] O. T. Seierstad, "Microsoft Windows CardSpace and the Identity Metasystem," *Telektronikk,* vol. 103, p. 9, 10/2009 2007.

[4] A. p. a. M. Administrativa. (2007, November 3th 2009). *Cartão de Cidadão (Portuguese Citizen Card)*. Available: http://www.cartaodecidadao.pt/

[5] E. Goffman, *The Presentation of Self in Everyday Life*: Anchor Books, 1959.

[6] M. R. Leary and R. M. Kowalski, "Impression management: A literature review and two-component model.," *Psychological Bulletin,* vol. 107, p. 13, 1990.

[7] D. Boyd, *et al.*, "Representations of Digital Identity," presented at the Conference on Computer Supported Cooperative Work, Chicago, U.S., 2004.

[8] Y. D. a. L. Xu. (2007, 10/2009). *Evolution of the World Wide Web. Part 1: Past and Present of WWW. Web 1.0: a world of newborns.* Available: http://www.deg.byu.edu/ding/WebEvolution/evolution-review.html#web-10

[9] L. X. Yihong Ding. (2007, 10/2009). *Evolution of the World Wide Web. Part 1: Past and Present of WWW. Web 2.0: a world of pre-school kids.* Available: http://www.deg.byu.edu/ding/WebEvolution/evolution-review.html#Web-20

[10] E. Bertino, "Digital identity Management and Protection," in *Google TechTalks*, ed: Google, 2008.

[11] F. I. Assad Moini and Azad M. Madni, "Leveraging Biometrics for User Authentication in Online Learning: A Systems Perspective," *Systems Journal, IEEE,* vol. 3, p. 5, 12/2009 2009.

[12] D. Thanh and I. Jørstad, "The Ambiguity of Identity," *Telektronikk,* vol. 103, p. 7, 2007.

[13] P. Forrest, "The Identity of Indiscernibles," in *Stanford Encyclopedia Of Philosophy*, E. N. Zalta, Ed., ed, 2009.

[14] H. Noonan, "Identity," in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., ed, 2009.

[15] M. Valcheva, "Playlistism: a means of identity expression and self-representation," p. 16, 2009.

[16] S. A. A. Rajon and S. A. Zaman. (2008, 11/2009). Implementantion of E-Governance: Only way to Build a Corruption-free Bancgladesh., 6. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4802970

[17] F. Pimenta, *et al.* (2010, 06/2010). GlobaliD: Federated Identity Provider Associated with National Citizen's Card. 6.

[18] J. V. Boettcher and A. Powell. (2002, 10/2009). Digital Certificates. 10. Available: http://www.cren.net/crenca/docs/syllabus.pdf

[19] J. V. Boettcher, *et al.* (2003, 10/2009). Digital Certificates: Coming of Age. 2. Available: http://net.educause.edu/ir/library/pdf/erm0317.pdf

[20] J.-D. Aussel, "Smart Cards and Digital Identity," *Telektronikk,* vol. 103, p. 13, 10/2009 2007.

[21] K. Cameron. (2005, 11/2009). The Laws of Identity 12. Available: http://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf

[22] D. V. Thuan, "Identity Management Demystified," *Telektronikk,* vol. 103, p. 8, 2007.

[23] R. A. d. C. Ferreira, "Privacy and Identity Selection," Engenharia de Electrónica e Telecomunicações Masters, Departamento de Electrónica, Telecomunicações e Telemática, University of Aveiro, Aveiro, 2008.

[24] L. J. Camp, "Identity in Digital Government," Kennedy School of Government, Cambridge2003.

[25] R. Semančik, "Revised World Wide Web Architecture," Phd. Phd, Faculty of Informatics and Information Technologies, Slovak University of Technology and Bratislava, 2008.

[26] SpendOnLife.com. (2009, 10/2009). *2009 Identity Theft Statistics*. Available: http://www.spendonlife.com/guide/2009-identity-theft-statistics

[27] M. Corporation. (2008, *Online Identity Theft: Changing the Game Protecting Personal Information on the Internet*. Available: http://download.microsoft.com/download/0/d/3/0d34ccfa-5498-4fab-bb32-16c881bafba7/Online%20ID%20Theft-%20Changing%20the%20Game.pdf

[28] T. G. Organization, "Confidence in the Information Society," *Flash EB,* vol. 250, p. 59, 2008.

[29] K. Delac and M. Grgic, "A survey of biometric recognition methods," presented at the 46th International Symposium electronics in Marine, Zadar, Croatia, 2004.

[30] O. f. E. C.-o. a. Development, "The Role Of Digital Identity Management in the Internet Economy: A primer for policy makers.," t. a. I. Directorate for Science, Ed., ed, 2009, p. 21.

[31] H. M. Tessem and K. R. Skaaraas, "Creating a security culture," *Telektronikk,* vol. 103, p. 8, 10/2009 2005.

[32] A. J. Elbirt. (2005, 20/06/2005) Who are you? How to protect against identity theft. *Technology and Society Magazine, IEEE*. 4.

[33] J. Brainard*, et al.*, "Fourth-Factor Authentication: Somebody You Know," presented at the Computer and Communications Security, 2006.

[34] A. Conklin*, et al.*, "Password-Based Authentication: A system Perspective," in *Hawaii International Conference on System Sciences*, Hawaii, 2004, p. 10.

[35] P. Wang*, et al.*, "Strengthening Password-based Authentication Protocols Against Online Dictionary Attacks," presented at the Applied Cryptography and Network Security, Third (2005), 2005.

[36] S. Chakrabarti and M. Singhal, "Password-Based Authentication: Preventing Dictionary Attacks," *Computer,* vol. 40, p. 7, 25/06/2007 2007.

[37] S. Networks. (003, 04/2010). Comparing Web Authentication Methods. 3. Available: http://www.sevannetworks.com/wpapers/WebAuthenticationMethods.pdf

[38] G. Security. 05/2010). *Biometrics*. Available: http://www.globalsecurity.org/security/systems/biometrics.htm

[39] J. Ortega-Garcia*, et al.* (2004, 08/2004) Authentication Gets personal with Biometrics. *Signal Processing Magazine, IEEE*. 13. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1276113

[40] S. Seno*, et al.*, "A network authentication system with multi-biometrics," presented at the 9th Asia-Pacific Conference on Communications 2003.

[41] W. A. a. L. Tien. (2003, 2010/05). *Biometrics: Who is watching you?* Available: http://www.eff.org/wp/biometrics-whos-watching-you

[42] R. W. Younglove, "Public Key infrastructure," *Computing & Control Engineering Journal* vol. 12, p. 3, 04/2001 2001.

[43] P. Gutmann, "PKI: it's not dead, just resting," *Computer,* vol. 35, p. 0, 08/2002 2002.

[44] A. Yasinsac and J. Childs, "Analyzing Internet security protocols," presented at the Proceedings Sixth IEEE International Symposium on High Assurance Systems Engineering. Special Topic: Impact of Networking, Boco Raton, FL, USA, 2001.

[45] S. Farrell. (2010, 12/2009) Why Didn't We Spot That? *Internet Computing, IEEE*. 4. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5370826

[46] J. C. a. C. Guo, "Online Detection and Prevention of Phishing Attacks (Invited Paper)," presented at the Communications and Networking in China, 2006. ChinaCom '06., Beijing, 2006.

[47] G. Goth, "Phishing attacks rising, but dollar losses down.," *Security & Privacy,* vol. 3, p. 1, 2005.

[48] A. Jøsang and S. Pope, "User Centric Identity Management," presented at the AusCERT - Asia Pacific Information Technology Security Conference Refereed R&D Stream, Gold Coast, Australia, 2005.

[49] CA. (2007, 10/2009). The bussiness value of Identity Federation. Available: http://www.comnews.com/WhitePaper_Library/Security/pdfs/CAfedbiz_drivers.pdf

[50] CA. (2007, 10/2009). Identity Federation: Concepts, Use Cases and Industry Standards. Available: http://images.vnunet.com/v7_static/itw/pdf/identity_federation_wp.pdf

[51] P. Gray. 10/2009). *Protecting Privacy and Security of Personal Information in the Global Electronic Marketplace*. Available: http://www.ftc.gov/bcp/icpw/comments/ico2.htm

[52] M. Gupta and R. Sharman, "Dimensions of Identity federation: A Case Study in Finacial Services," *Journal of Information Assurance and Security 3,* p. 13, 2008.

[53] O. Foundation. 22th October 2010). *http://openid.net/get-an-openid/what-is-openid/*.

[54] O. Foundation. 10/2009). *Surprise You may already have an OpenID*. Available: http://openid.net/get-an-openid/

[55] O. Foundation. (2007, 10/2009). *OpenID Authentication 2.0 - Final*. Available: http://openid.net/specs/openid-authentication-2_0.html

[56] H.-K. Oh and S.-H. Jin, "The Security Limitations of SSO in OpenID," in *The 10th International Conference on Advanced Communication Technology*, Gangwon-Do, South Korea, 2008, p. 4.

[57] M. Corporation. (2006, 23th October 2009). *Introduction to Windows Live ID*. Available: http://msdn.microsoft.com/en-us/library/bb288408.aspx

[58] K. Young. (2009, 11/2009). *Windows Live ID Identity gateway for Microsoft online services*. Available: http://winliveid.spaces.live.com/

[59] C. J. V. Teixeira, "Infra-estrutura para portal internet integrador de serviços," PhD PhD, Departamento de Electrónica , Telecomunicações e Informática, University of Aveiro, Aveiro, 2009.

[60] T. L. A. Project. (2001, 10/2009). *About*. Available: http://www.projectliberty.org/liberty/about/

[61] T. L. Alliance, "Specifications," March of 2005 2005.

[62] O. Standards, "Oasis Web Services Security (WSS)," 2006.

[63] Internet2. (2001, 11/2009). *Shibboleth - Specification - DRAFT v1.0*. Available: http://shibboleth.internet2.edu/docs/draft-internet2-shibboleth-specification-00.html

[64] M. B. Jones, "The Identity Metasystem: A User-centric, Inclusive Web Authentication Solution," presented at the W3C Workshop on Transparency and Usability of Web Authentication, New York City, 2006.

[65] O. Standards. (2005, 11/2009). *Saml Specifications*. Available: http://saml.xml.org/saml-specifications

[66] OASIS. 11/2009). *Organization for the Advancement of Structured Information Standards*. Available: http://www.oasis-open.org/who/

[67] OASIS, "Assertions and Protocols for the OASIS Security Assertion Markup Language (SAML) V2.0 – Errata Composite," ed, 2009, p. 91.

[68] OASIS, "Bindings for the OASIS Security Assertion Markup Language (SAML) V2.0 – Errata Composite," ed, 2009, p. 46.

[69] OASIS, "Profiles for the OASIS Security Assertion Markup Language (SAML) V2.0 – Errata Composite," ed, 2009, p. 69.

[70] OASIS, "Metadata for the OASIS Security Assertion Markup Language (SAML) V2.0 – Errata Composite," ed, 2009, p. 45.

[71] OASIS, "Authentication Context for the OASIS Security Assertion Markup Language (SAML) V2.0," ed, 2005, p. 70.

[72] OASIS, "Security and Privacy Considerations for the OASIS Security Assertion Markup Language (SAML) V2.0," ed, 2005, p. 33.

[73] OASIS, "Conformance Requirements for the OASIS Security Assertion Markup Language (SAML) V2.0," ed, 2005, p. 19.

[74] I. Jørstad, *et al.*, "Strong Authentication for Internet Applications with the GSM SIM," *Telektronikk*, vol. 103, p. 9, 2007.