

Alexandre Barbosa Augusto

**A Mobile Based Attribute
Aggregation Architecture for
User-Centric Identity Management**



Departamento de Ciência de Computadores
Faculdade de Ciências da Universidade do Porto
Setembro de 2012

Alexandre Barbosa Augusto

A Mobile Based Attribute Aggregation Architecture for User-Centric Identity Management

*Dissertação submetida à Faculdade de Ciências da
Universidade do Porto como parte dos requisitos para a obtenção do grau de
Mestre em Engenharia de Redes em Sistemas Informáticos*

Orientador: Prof. Manuel Eduardo Carvalho Duarte Correia,
Professor Auxiliar do Departamento de Ciência de Computadores
da Faculdade de Ciências da Universidade do Porto

Departamento de Ciência de Computadores
Faculdade de Ciências da Universidade do Porto
Setembro de 2012

I have to thank my advisor, Manuel Correia, for believing and supporting me, my family for all its support, my friends for being there always when i needed them, Casa Melo for the memories (including the ones which i do not remember) and, last but not least, my love, Cátia, for supporting me and for not killing me when i was (rarely) sleeping during the thesis process.

Alexandre Barbosa Augusto
September 2012

Acknowledgments

I have to thank my advisor for his unconditional support during all the process of OFELIA project that led the way for this thesis work.

I also have to thank the ERDF through the Programme COMPETE and by the Portuguese Government through FCT - Foundation for Science and Technology, the project OFELIA ref. PTDC/EIA-EIA/104328/2008 and is being conducted with the institutional support provided by DCC/FCUP and the facilities and research environment gracefully provided by the CRACS (Center for Research in Advanced Computing Systems) research unit, an INESC TEC associate of the Faculty of Science, University of Porto.

Resumo

O crescimento explosivo da Internet está a acelerar a migração das infraestruturas sociais essenciais do mundo real para o mundo virtual. A gestão de identidade tem desempenhado um serviço fundamental neste processo, tornando mais fácil planear, distribuir e proteger o acesso online. Recentemente o desenvolvimento em modelos de gestão de identidade estruturados à volta de conceitos centrados no utilizador tem sido foco de atenção. Este processo está assim de acordo com a exigente sociedade digital e empoderamento do utilizador com ferramentas digitais que permitem uma gestão da identidade digital centrada no utilizador de uma forma mais fiável, responsável e segura. Incidentes recentes ligados com a divulgação não autorizada de informação sensível, mostram o quão importante é para o utilizador conseguir exercer algum controlo sobre aquilo que sobre ele é publicamente conhecido e disseminado na Internet. Deste modo é imperativo promover o desenvolvimento de sistemas interoperáveis padronizados que permitam uma gestão focada no utilizador de informação privada e ajudar a proteger os direitos fundamentais de privacidade do utilizador.

Fornecedores de identidade tal como Google e Facebook, estão atualmente numa feroz competição por utilizadores. Um dos seus principais objetivos é a criação de enormes e monopolizadas bases de dados pessoais, permitindo a produção de perfis de utilizadores altamente precisos que podem ser monetizados eficientemente para fins de marketing e "lock-in". Alguns tipos de dados pela sua natureza podem ser sujeitos a constantes mudanças ficando assim facilmente desatualizados. Com um gestor de identidade centralizado e "distante", pode tornar-se difícil gerir o grau de atualização para esses dados pessoais dinâmicos. Por todas estas razões, acreditamos que toda a informação privada do utilizador deve ser mantida o mais próximo possível da fonte primária do seu proprietário e sobre seu controlo direto. Estes dados podem então ser geridos pelo utilizador com a ajuda de entidades conhecidas como Autoridade de Atributos (*Attribute Authority* - AA) pessoais na rede. Estes AAs partilham dados pessoais a outras aplicações apenas depois do consentimento explícito do proprietário que deve ter em

consideração a identidade do Provedor de Serviços (*Service Provider/Relying Party*). Este consentimento deve assumir a forma de autorização condicional, temporalmente limitada e facilmente revogável. Estas condições são essenciais para a gestão focada no utilizador, que constitui a base da nossa proposta para desenvolvimento de um modelo totalmente descentralizado, focado na privacidade e no utilizador para gestão de identidade baseado na agregação de dados privados distribuídos e protegidos por um conjunto de AAs pessoais.

Abstract

The explosive growth of the Internet is accelerating the translation of essential real world societal infrastructures to the virtual world. Identity management has been playing a fundamental structuring service role in this process, by making it easier to plan, deploy and secure online access to an ever increasing number of systems and applications. More recently, the general tendency has been to concentrate development efforts on identity management models structured around user-centric concepts, totally in concert with a digital society that is demanding and is ever more focused on empowering individuals with tools for a more reliable, responsible and secure user-centric management of private digital data. Recent incidents related with the unauthorized disclosure of sensitive information also shows how important it is for users to be able to exercise some control on how much about them is publicly known and disseminated on the Internet. It is therefore crucial to promote the development of standardised interoperable systems that enable the user-centric management of private information and help secure users basic right for privacy.

Massive centralized identity providers (Google, Facebook), are currently under a fierce competition over the hearts and minds of users. One of their main purposes is to create enormous monopolized user personal data database as they enable them to produce highly accurate user profiles that they can then monetize very efficiently for marketing and lock-in purposes. There are also some types of important personal data that by their very nature can be subjected to change and thus become stale, sometimes very quickly. With a centralized and distant identity provider it can therefore become quite difficult to manage the degree of staleness for dynamic personal data. For all these reasons we believe that all users private information should be kept as much as close as possible to their owners primary source, under his direct control. This data can then be managed by the user with the help of personal Attribute Authority(AA) entities in the network. These AAs disclose personal data to other applications only after the owners explicit consent, which should also take into

consideration the identity of the original requester at the Service Provider/Relying Party (SP/RP). These consents should also take the form of conditional, temporal limited and easily revocable authorizations. These essential general assumptions for user-centric management of private data constitute the base for our proposal on the development of a fully decentralized privacy and user-centric model for identity management based on the aggregation of users private data, distributed and protected by a set of personal AAs.

Palavras Chave / Keywords

Palavras Chave:

- Agente de autorização
- Centrado no utilizador
- Identidade móvel
- Agregação de atributos
- Controlo de acesso
- Dispositivos móveis
- Gestão de identidade
- Identidade digital

Keywords:

- Authorization broker
- User centric
- Mobile Identity
- Attribute aggregation
- Access control
- Mobile device
- Identity management
- Digital Identity

Acronyms

CA certification authority

HTML Hypertext Markup Language

HTTP hypertext protocol

IdM Identity Management

IdMS Identity Management Systems

IdP identity provider

IETF Internet Engineering Task Force

IP internet protocol

PKI Public Key Infrastructure

SAML Security Assertion Markup Language

SSO Single Sign-On

STS Security Token Service

TLS Transport Layer Security

SSL Secure Socket Layer

XML Extensible Markup Language

AA Attribute Authority

SP Service Provider same as RP

RP Rely party same as SP

PGP Pretty Good Privacy

GPS Global Positioning System

EER Enhanced Entity Relationship

JDK Java Development Kit

Contents

Acknowledgments	vii
Resumo	ix
Abstract	xi
Palavras Chave / Keywords	xiii
Acronyms	xv
List of Tables	xxiii
List of Figures	xxvi
1 Introduction	1
1.1 Motivation	3
1.2 Proposed Solution	4
1.2.1 Objectives	5
1.2.2 Features	5
1.2.3 Applications	6
1.3 Application Domain	7
1.3.1 Identity Management	7

1.3.2	eHealth	8
1.3.2.1	Patients EHR (Access and Control)	8
1.3.3	Physical Security Access	9
1.4	Contributions	10
1.5	Outline	12
2	State of the Art	13
2.1	Research Methods	13
2.1.1	Identity management systems	15
2.1.2	Traditional model	16
2.1.3	Federated Model	17
2.1.3.1	Shibboleth	18
2.1.4	Centralized Model	19
2.1.4.1	OpenID	20
2.1.4.2	Microsoft .NET Passport	21
2.1.5	User Centric Models	21
2.1.5.1	Higgins	21
2.2	Attribute Aggregation	22
2.2.1	Client mediated assertion	23
2.2.2	Identity Federation model	23
2.2.3	Relying Party mediated attribute aggregation	23
2.2.4	Identity proxying/chaining	24
2.2.5	Identity Relay	24
2.2.6	Linking Service	24
2.3	Identity Management Common Standards	25
2.3.1	Kerberos	25

2.3.2	Security Assertion Markup Language (SAML)	26
2.3.3	Extensible Access Control Markup Language (XACML)	27
2.3.4	Web Services Federation (WS-Federation)	27
2.3.5	Digital ID wallet	28
2.3.6	Open Identity Exchange (OIX)	28
2.4	Security Concepts	29
2.4.1	Public Key Cryptography	29
2.4.1.1	Public Key Infrastructure	29
2.4.1.2	Digital Certificates	30
2.4.2	Smartcard	33
2.4.2.1	Mobile Secure Card	33
2.4.3	Valet key based protocol	34
2.4.3.1	Open Authorization (OAuth)	34
2.4.3.2	Azure Microsoft	35
2.5	Communication Concepts	35
2.5.1	XMPP	35
2.5.2	Web service	36
2.5.2.1	REpresentation State Transfer (REST)	36
2.5.2.2	Simple Object Access Protocol (SOAP)	37
2.5.3	Quick Response code	38
3	Architecture Components	39
3.1	Architecture Technologies Overview	39
3.1.1	microSD mobile card	39
3.1.2	The XMPP messaging protocol	40
3.1.3	Secure access authorization tokens	41

3.1.4	The OFELIA TRUST infrastructure	42
3.1.5	XML Schema	42
3.1.6	QR Code	44
3.2	Architecture nodes overview	44
3.2.1	OpenID Connect Identity Services	44
3.2.2	Relying Party/Service Provider (RP/SP)	44
3.2.3	Attribute Authorities	45
3.2.4	The Identity Broker	46
3.2.5	The smartphone as a Secure Digital Wallet	46
4	Entities enrollment and case scenario	49
4.1	Attribute Authority enrollment	50
4.2	Identity broker enrollment	51
4.3	Service provider enrollment	53
4.4	Usage case Scenario	55
4.4.1	Attribute Aggregation scenario in e-commerce	55
4.4.2	Attribute authorization access scenario in healthcare	56
5	Conclusions and Future Work	59
5.1	Research summary	59
5.2	Main findings	60
5.3	Current OFELIA implementation limitations	61
5.4	Future work	62
5.5	Conclusion	62
A	Development notes	65
A.1	Identity Broker Web Service	65

A.2 Attribute Authority standalone application	68
A.3 Secure Digital Wallet application	70
A.4 Relying Party web service	71
A.5 Screen captures:	74
References	76

List of Tables

2.1	The Seven Laws of Identity	15
2.2	Traditional model limitations	16
2.3	Information Security Attributes	25
2.4	HTTP verbs and their meaning in RESTful web servers	37

List of Figures

2.1	Federated model example	17
2.2	Single sign-on model example	19
2.3	Smartphone user centric model	22
2.4	Components of the SAML Specifications	26
2.5	Public key cryptographic signature process	30
2.6	Public key cryptographic encryption process	31
2.7	Certification validation by the usage of a Certification Authority	32
3.1	The purposed architecture communication	40
3.2	Architecture data exchange XML Schema	43
4.1	AA enrollment flow	51
4.2	Identity Broker enrollment flow	52
4.3	RP/SP enrollment flow	54
4.4	Agnes access permissions to Katherines' breast pathology folder	57
A.1	IdB package tree	67
A.2	Identity Broker Database EER	68
A.3	AA Android package tree	69
A.4	Attribute Authority Database EER	70
A.5	Digital wallet Android package tree	71

A.6	Secure digital wallet Database EER	72
A.7	Relying party package tree	72
A.8	Relying Party Database EER	73
A.9	Ofelia digital wallet authorization request box.	74
A.10	Ofelia digital wallet token list.	74
A.11	Relying Party user area.	74
A.12	Relying Party data request page.	75
A.13	Relying Party GPS coordinate consultation.	75

Chapter 1

Introduction

There is a massive privacy problem on the Internet that needs to be solved in a more meaningful and user centric way. The problem lies not only on the lack of trust but also on how users can effectively regain control by exercising privacy rights [1] in a more meaningful way. There are many users that spend a very significant part of their lives online, thus creating and many times unconsciously contributing with enormous quantities of personal data to less scrupulous service providers, which can then easily become as individually traceable as the user's own biological DNA. And like the DNA, this data can be used, and it is in fact been abused by service providers, to track and store private personal information that users usually would not want to have publicized. Controlling our own data and ensure its privacy is the only way to secure and at the same time maintain the intrinsic value of our personal information preventing it from being trivialized, exploited or shared without our consent or have a rightful share of the profits associated with the exploitation of its value. It is thus vitally important to determine and set a hard limit on what should be deemed acceptable service provider behaviour and what constitutes a flagrant violation of our constitutional rights, after all it is our own digital personal data that is at stake.

All this is happening right under our noses. The massive growth of the Internet is being fed by an ever increasing set of apparently free online services that are being sustained by the still highly unregulated mercantilization of the users personal data. Remember that when a popular Internet service is "free" and it is not trying to sell you anything it generally means that it is the user that is the product being sold. Traditionally, many web sites require the user to disclose hitherto unnecessary sensitive identity information, like postal addresses, telephone numbers, gender, credit cards, geographic location and even otherwise highly sensitive personal data like for example health

records (Google Health, Microsoft Healthvault) in order to provide an apparent "free" service. In essence these services are definitely not free and constitute a means by which the user is being mercantiled by the service provider. As an example consider the current fierce competition being fought off by Google and Facebook about digital identity services. This has led to the development of massive identity infrastructures that accelerated the deployment of efficient and easily deployable identity and access control protocols like OpenID [2] and OAuth [3] to better cater the needs of the service providers web applications. Unbeknownst to users they are currently being played as pawns in a very complex scenario that they do not fully understand, where the real value of their identity attributes is not disclosed and where their identity and privacy is constantly being negotiated between a set of internet services, without their consent or opportunity for control [4].

In order to better contextualize and understand the relevant real issues behind this problem we have conducted an highly comprehensive research and analysis of the published literature on *identity management, user centric access control mechanisms and identity attributes aggregation and management models*. This led to the development of a novel user centric framework architecture for identity attributes aggregation (OFELIA - Open Federated Environment for the Leveraging of Identity and Authentication) that has been designed right from the start to cater with the problems related with the service providers abuse of confidence over matters related with digital identity management. Our architecture defines an innovative user centric authorization/authentication mechanism that protects users' privacy by the means of novel discretionary access control mechanisms, that are managed and exercised by the means of mobile devices, more specifically the users smartphones. The smartphone nowadays provides tremendous opportunities because it can also be seen as a truly personal platform for sensing the users surrounding environment. They are provided with cameras, microphone, gyroscope, GPS and a plethora of other sensing equipment that can be connected via bluetooth, like for example the cardiac rhythmic sensors that are nowadays so popular within the jogging community. To cater with this so much richer universe of highly dynamic personal data, with OFELIA we have also expanded the set of static identity attributes, being currently managed by standard Internet identity management systems [5, 6], with a new set of highly changeable identity attributes. This opens a whole new range of opportunities and possibilities by having real time dynamic personal data being *Processed As Requested* (PaR). In other words, every time a Relying Party (RP) or Service Provider (SP) requests a dynamic mutable identity attribute, data retrieval occurs in real time and

goes as deep as its authoritative origin. This new digital identity paradigm is useful in contexts where identity attributes are highly volatile, because of its high rate of change, where PaR becomes mandatory if the service provider objective is to keep a coherent timeline for the identity attributes being thus retrieved. This is easily illustrated with for example GPS positioning. If we consider applications where this is an identity attribute (the users current location), in situations where the users are in constant movement, the more recent the value for this attribute is the more valuable is for the service provider. For example service providers (relying parties) would be very keen to serve high value banner advertisements based on the user current user's GPS positioning. The more current, the more valuable the advertisement would be. This can only happen within a PaR paradigm where the Relying Party is provided with a digital identity infrastructure where it can obtain the real current time positioning of an individual and not the last time the user or application remembered to update it. This of course opens a new Pandora Box of privacy violation and surveillance nightmares that can only be coped if user is given full control over who can access, for how much time and on what conditions can this highly dynamic real time data be retrieved from the source by the service providers.

1.1 Motivation

After graduating in Computer Engineering Sciences at faculty of Sciences of University of Porto, the author started his master degree in Network and Information Systems Engineering following the Communication Networks minor at the same institution of his graduation. The first year of his master degree amplified the authors' interest in computer security, leading the author to develop a special taste over the complexity of digital identity management.

During the second year the author joined the OFELIA (Open Federated Environment Leveraging Identity and Authorization, ref. PTDC/EIA-EIA/104328/2008) project, that acted as the main motivation source for this thesis. This work was funded by the ERDF through the Programme COMPETE and by the Portuguese Government through FCT - Foundation for Science and Technology. The main institutions behind the project are the faculty of science (CRACS research unit an INESC TEC associate), faculty of medicine (CINTESIS research unit) both from the University of Porto, the University of Minho and the Portugal OpenID foundation. The purpose of this project is to research the conceptualization and implementation of new ideas and extensions

for user centric identity management and authorization mechanisms for the Internet, where sensitive identity attributes can be directly stored on the users personal mobile devices or other user controlled secure storage services on the Internet, whose location can be hidden from the service providers to enhance and protect users privacy.

Our first identity management framework model, implemented by the author for the OFELIA project, allowed users to asynchronously control and effectively share sensitive dynamic identity attributes directly held on mobile devices (in this case android smartphones) [7]. This initial model served as the starting point for the rather more complex process of effectively decoupling the storage of identity attributes from the mobile phone, and it is this rather more complex identity management infrastructure that is described in this thesis.

1.2 Proposed Solution

We have developed an identity/access control model that aims the development of standardised interoperable systems that enable the user-centric management of private information by the usage of smartphones. This model establish the right means to the users to guarantee the basic right for privacy over their valuable identity attributes. The adoption of smartphones as a user-centric management platform is highly appropriate because these devices are nowadays ubiquitous, have more then adequate processing CPU power to run modern operating systems, are being deployed with full Internet access, are accompanied by fully matured development systems and constantly follow their owners everywhere as the "*de facto*" personal mobile device. Therefore providing a practical solution for the users Internet reachability challenge [8]. The use of smartphones for identity management is currently also recognized as essential for enhancing security and privacy [9, 10, 11] and has been proved to play a crucial role on more flexible user-centric models [12, 13]. Based on these facts, the proposed architecture adopt the use of smartphones as an Authorization Broker in order to grant users a dynamic and more active role over their identity attributes.

The proposed architecture has been divided in four different components: (1) one application programming interface (API) in order to establish a fast deployment of the architecture for the relying parties and service providers; (2) other API for the Attribute Authorities (AAs) creating the necessary security means to allow data access; (3) an implementation of an identity broker based on the linking service aggregation model referred on section 2.2.6; (4) and an android application, implementing the

secure mobile authorization broker to aggregate AAs and authorize, manage and revoke access to their identity attributes over the different AAs.

1.2.1 Objectives

The main goals of this thesis is to define and implement a user-centric distributed architecture for identity management by the means of mobile devices. We have identified and proposed the following main objectives:

- **Research** of the current state of the art on identity management and attribute aggregation.
- **Definition** of an architecture that creates an authentication/authorization mechanism to protect users privacy by using the research results.
- **Implementation** of the defined architecture that can be used to fulfill one of the main goals of the OFELIA project.
- **Deployment** of the implemented architecture in real scenarios in order to analysis the architecture performance and usage.
- **Dissemination** of the results on identity management and access control related symposiums and international conferences in order to diffuse our results thus enriching our research.

1.2.2 Features

The main features of the proposed architecture are:

- **Attribute Aggregation** - Establish the right means for a user aggregate their personal attributes and at the same time keep in the control over the data these characteristics are archived by the usage of the identity broker and the authorization broker.
- **User centric model** - Personal attributes are only disclosure with previous user authorization. The owner of personal attributes maintain the capacity to revoke or limit the access at any time.

- **Dynamic Attributes** - These type of attributes are temporal sensitive and can only reside, not in the cloud, but in mobile personal devices in order to keep those values up to date creating the necessary temporal stream.
- **Easy Integration** - The architecture provides two distinct application programming interface in order to allow all kind of relying parties and attribute authorities quickly integrate in the proposed architecture.
- **High privacy level** - Personal attributes are only disclosure for an authorized requester. Despite the fact that every personal attribute pass through the identity broker the data exchanged is always encrypt by the requester certificate and sign by the data attribute authority.
- **Decentralized solution** - The Users' digital information are fragmented through various attribute authorities, decreasing the chance of success attacks in order to obtain a user full digital information.

1.2.3 Applications

In the course of our work we have developed several applications to further illustrate and explore the attribute aggregation architecture for user centric identity management that is the objective of this thesis. They are:

1. **A mobile application:** that implements a secure mobile authorization broker to aggregate attribute authorities and authorize, manage and revoke access to their identity attributes allowing the user to asynchronously exercise discretionary access control over its identity attributes in a simple and highly transparent way.
2. **An application programming interface for relying parties:** in order to establish a fast deployment of our proposed architecture in the most diverse types of relying parties.
3. **An application programming interface for attribute authorities:** in order to create the necessary security means to allow data access.
4. **A standalone application:** that implements a secure attribute authorities. This application was developed in Java language in order to provide system portability.

5. **Identity Broker:** a web service that acts as an identity proxy, based on the linking service aggregation model.
6. **A QR Registration plugin application:** in order to establish a fast account creation between the mobile device and XMPP server.

1.3 Application Domain

Identity management system has become an essential tool for economical evolution on almost every sector. This section further explores some sectors that show great potential for the deployment of our proposed architecture.

1.3.1 Identity Management

The interest on user digital identity has been increasing dramatically over the recent years due to the discovery of its highly strategic commercial value for the market [14]. Massive centralized identity providers started to flourish, companies like Google, Facebook and even Microsoft, are currently under a fierce competition over the hearts and minds of users for their personal data. One of their main purposes is to create enormous monopolized centralized databases of user identity attributes as they allow them to produce highly accurate user profiles that they can then monetize very efficiently for marketing and lock-in purposes [15]. These global companies harvest and aggregate personal data in such a large scale that, lest it is put under some kind of control, it will very soon represent a major global threat to personal security and privacy the like of which the world has never seen.

Due to this competition over digital identity, protocols like OpenID for authentication and OAuth [16] for authorization started to appear in order to cater the need for web applications data interoperability creating the necessary means for centralized identity providers. These protocols are employed as standardized mechanisms to build single sign-on systems and attribute sharing based on valet keys [17]. More recently a new open standard named as OpenID Connect [18] is under development, as a single solution for aggregating both authentication and authorization.

In the middle of this dispute lies the user that usually do not understand where his identity and privacy is being negotiated between a set of fancy internet services and sometimes even without his consent or control [4]. This privacy abuse is not the only

problem, if a centralized identity provider like Google suffers an attack, millions of highly well detailed personal attributes will be compromised.

In order to struggle against these massive centralized identity providers our proposed architecture defined that the users private information should be kept as much as close as possible to their owners primary source, under his direct control and never with a "trusted" third party like *Facebook* for instance. This data should be managed by the user with the help of personal Attribute Authority (AA) entities in the network. These AAs disclose the personal digital data to Relying parties or Service Providers only after the data owner authorize it. These authorization take the form of conditional, temporal limited and easily revocable by the data owners. [7, 13]

1.3.2 eHealth

In the healthcare domain, patients digital data is normally collected into what is called the Electronic Health Record (EHR). The EHR encompasses many functions that can include different types of data items such as diagnoses, medications and operations [19, 20]. The EHR is nowadays indispensable for health institutional purposes and could be used to empower patients by giving them the necessary information to play a more active role in their own health and in their families health as well [21].

OFELIA has the potential to change current eHealth practices and business practices by giving the patients the means by which they can directly manage and exercise revocable access control to their EHR and thus effectively decide with whom to share their EHR, be it healthcare professionals or even family members and caring friends [22].

1.3.2.1 Patients EHR (Access and Control)

Nowadays, patients want to be better informed about their medical conditions and play a more active role on their own treatments. They usually consult online information by using search engines to educate themselves about etiology, treatments, and the prognosis of the medical conditions. Getting access to their medical records would help the patient to better understand their medical conditions [23]. On this issue the European Recommendation [24] and American Legislation [25] for protection of medical data agree that the patient must have access to his/her medical record and play a major role in the decisions regarding the content and the distribution of his/her

medical data [26].

Currently, for patients to have access to their medical records they need to write a request to their custodian healthcare institution and the response delay depends on their country legislation (e.g. 10 days in Portugal [27], 21 days in England [28] and 30 days in USA [25]). We believe that the latest developments in digital communications and information technologies should provide the patients a simpler and more secure way to access their medical records and at the same time provide a better collaboration and interaction experiences with the healthcare professionals [29].

Unfortunately, by opening up the access to the EHR with inadequate access control mechanisms and policies carries some substantial risks as illustrated by the grim statistics observed during the period of 2006 to 2007, where in the USA, over 1.5 million names were exposed during data breaches that occurred in hospitals [30]. One of the most important and complex requirement for eHealth systems [31] is to keep patients' information private and secure and that is exactly where our proposed framework could be framed, acting as the authorization infrastructure that acts as node between the health institution patient records and the patient

Patient medical records exchange between health institutions is very difficult due to problems ranging from incompatible platforms and data formats, passing through difficult to negotiate access policies that are acceptable by all participating Institutions, to the lack of a common agreed upon identity/authorization federated infrastructure for both patients and health professionals that would allow them to assume compatible/recognizable roles within each hospital eHealth system. Under this very difficult scenario, our proposed framework, can again act as the authorization infrastructure that allows the patient to share their medical information with health professionals independently of the platform since the health professionals would be accessing the desired data directly from the data source. This scenario would create a false but helpful kind of interoperability between health institutions

1.3.3 Physical Security Access

The key is the most used token of all time to give/have access to a house, a car, a safety box and etc. The most common keys are the house keys and the car keys. Usually the usage of these keys have several problems like:

- Easily duplicable, most of the keys are easy to copy.

- Non revocable, once you give it you have to change your lock in order to "revoke" the access.
- Its a physical device, if you take it by mistake you can only solve this problem by returning it or having a copy.
- Not possible to define special conditions in the key usage, in other words its not possible to set usage polices.

By the usage of the proposed architecture we aim the use of smartphones as master devices that could be used as a key to start a car or/and open a house. This master device could also virtually authorize other devices to have access by setting special policies access like temporal constrains. This scenario would help to overcome the presented problems of the traditional key usage since the master device could manage all authorizations by limiting its access and revoking it at any time. Obviously the usage of mobile devices as keys to physical access will create a set of new issues related with security, however this issues could also be overcome with research.

1.4 Contributions

This section describes the contributions that were obtained during this research. These include: articles, book chapters, journals and technical reports that were published or submitted waiting for approval and as well as the studies and prototypes developed in this thesis process.

Articles in Journals

- A paper in Computers in Biology and Medicine journal titled as "A proposal for a secure patient empowerment architecture" from Elsevier. (submitted and waiting for approval)

Book Chapters

- A book chapter on the book named "Innovations in XML Applications and Meta-data Management: Advancing Technologies", titled as "A secure and dynamic mobile identity wallet authorization architecture based on a XMPP messaging infrastructure" from IGI Global.

- A book chapter on the book named "Architectures and Protocols for Secure Information Technology", titled as "A Mobile Based Attribute Aggregation Architecture for User-Centric Identity Management" from IGI Global. (accepted and waiting to be published)

Articles in Conferences

- A paper in International Information Security and Privacy conference (IFIP SEC 2012) titled as "OFELIA - A Secure Mobile Attribute Aggregation Infrastructure for User-Centric Identity Management" held in Crete, Greece.
- A paper in Information Technology in Bio- and Medical Informatics conference (DEXA, ITBAM 2012) titled as "A Mobile Based Authorization Mechanism for Patient Managed Role Based Access Control" held in Vienna, Austria.
- A paper in Conferencia Iberica de Sistemas y Tecnologas de Informacins (CISTI 2012) titled as "A literature review of security mechanisms employed by mobile agents" held in Madrid, Spain
- A paper in XML: Aplicaes e Tecnologias Associadas conference (XATA 2011) titled as "An XMPP messaging infrastructure for a mobile held security identity wallet of personal and private dynamic identity attributes", held in Vila do Conde, Portugal. (best paper award)

Studies and Prototypes:

- A comprehensive research over identity management models including the aggregation of attributes.
- The designing of a simplified identity semantic language in xml schema.
- The design and development and specifications of a user-centric distributed infrastructure for identity management by the means of mobile devices.

Technical Reports

- A technical report in Center for Research in Advanced Computing Systems (CRACS) titled as "OFELIA: Open Federated Environment for the Leveraging of Identity and Authorisation".

International conferences administrative work

- Member of the organizer committee of the special track on: Security and Privacy in Healthcare IT in the 26th IEEE International Symposium on Computer-Based Medical Systems (CBMS2013).

1.5 Outline

The remaining chapters of this thesis are organised as it follows:

Chapter 2 Presents an overview how digital identity is managed by the different models of identity management systems and how the mainly attributes aggregation models work. It finishes by explaining the methods to establish and secure a communication channel.

Chapter 3 Describes the proposed architecture by explaining each node of the proposed architecture and contextualizing with the used technologies and methods.

Chapter 4 Describes in detail the protocol messages exchanged to establish the connection between the architecture components and describes a usage case scenario, which can be quite useful to help to better understand the different components interaction.

Chapter 5 Discusses the presented work and the future directions.

Chapter 2

State of the Art

The explosive growth of the Internet is being supported by the translation of essential real world societal and monetary infrastructures to the virtual world, where digital identity plays a central catalyzing role that is accelerating the whole process. A digital identity can also be seen as being composed by a set of personal data attributes that in some sense characterizes one of the many roles a real person may assume in the virtual world when he uses an application. Like in the real world, where different persons can assume different roles depending on the situation and context, in the virtual world, digital identity is the process by which one can manage a set of personal attributes that are appropriate to assume a contextual identity, usually referred as an identity persona[32]. The association between a persona and a user is achieved by the means of an authentication process that is managed by an identity management system [33]. A successful authentication assures that a certain entity has the right to a certain set of personal credentials that the information systems can then use to determine the kind of resources that persona has the right to use. In this chapter we review the identity management models currently accepted and described in the literature, including identity attribute aggregation models and the necessary security and communication concepts and technologies necessary to properly understand them.

2.1 Research Methods

The literature review we have conducted to better contextualise the research and work we have done for this thesis was performed on January 2012, using the IEEE Xplore, ACM Digital Library and Google Scholar search engines. We have applied the

following queries:

- "[ANY FIELD] (((((((user centric model) AND security) AND privacy) OR identity management) AND digital identity) OR) attribute aggregation) OR Federated model)" in IEEE Xplore
- [ANY FIELD] identity management OR attribute aggregation OR user centric OR Federated identity management in Google Scholar
- "[ANY FIELD] (((((((user centric model) AND security) AND privacy) OR identity management) AND digital identity) OR) attribute aggregation) OR Federated model)" in ACM Digital Library

Due to the high number of results thus obtained we had them filtered by their abstracts, according to the following inclusion criteria:

1. English as language;
2. Recent articles, less than five years old;
3. Articles with high relevance on the search engines;
4. Well known and highly referenced authors in the field of digital identity management;
5. A cautiously review of titles, abstracts, keywords and section titles, to exclude papers with the same or somewhat outdated subjects.

After the analysing the most promising papers/standards texts, their citations were also reviewed and those that suited the inclusion criteria were also integrated in the review. The search and full text retrieval of the selected papers was performed in the following databases:

- Biblioteca do Conhecimento Online (b-on),
- Open Repository at the University of Porto,
- Open Access Scientific Repository of Portugal (RCAAP),
- Google Scholar

In the end we have selected 26 papers. Due to the advisor recommendation, we have also included six extra papers, amounting to 32 papers that were fully reviewed to determine a more accurate and useful context for the work we have done for this thesis.

2.1.1 Identity management systems

Digital identity is maintained by identity management systems that are systems composed of policies, economic model, business processes and technologies that implement and manage the personal identity users attributes (personas) needed to establish and manage access rights to the organization digital resources [34]. Moreover Identity management systems are also responsible for the digital identity lifecycle management within organizations as they provide the infrastructure deployed to validate and exchange the digital personal data attributes needed to establish and promote interoperability among different systems in accordance with some set of security and legal policies. According to Kim Cameron, to be useful, an Identity Management System (IdMS), must follow the “*seven Laws of Identity*” [35]. (see table 2.1) Identity

Table 2.1: The Seven Laws of Identity

	Description	Comments
One	User consent	An identity is identified and used only when the user agrees to it.
Two	Limited disclosure	The system provides the minimum identifying information required for the transaction.
Tree	Fewest parties	Only relying parties that need to know receive identifying information.
Four	Directional identity	Omni-directional versus uni-directional
Five	IdMS should work with a variety of identity technologies, run by multiple providers.	Designers cannot assume the feasibility of a universal identity or the availability of a single expression of an identity.
Six	Human integration	High levels of reliability between the human user and the system
Seven	Consistent experience across platforms	Similar to the way the web appears to users

management systems are employed by identity providers [33] that can assume several distinct models depending on the social and/or financial benefits.

2.1.2 Traditional model

The term identity management indicates the need to verify the identity of the entity accessing the system and the negotiation of rights and privileges based on privacy requirements and identity attribute disclosure associated with that identity. Identity management architectures could be server-based, client-based, or networked-based. Most common architectures are proprietary, application driven, and server-based. E.g. Amazon maintains its own users information, resulting in a silo approach, which is the most common identity management approach currently employed by the vast majority of Internet service providers. With this approach the user ends up being forced to have too many digital identities, each one of them managed by a different entity and requiring its own resources.

Moreover today's mainstream identity management in the Internet does not align well with most users, since it requires too much effort from the users (they have to maintain a huge collection of user credentials, one for each site they use) and a lot of resources for service providers (resources allocated specifically for identity management that could have been easily shared with other service providers lest they employed other more versatile identity management systems). In addition, a Silo based approach is not compatible with web services, ad-hoc and mobile computing, and possesses a plethora of well know limitations and liabilities. (see table 2.2)

Table 2.2: Traditional model limitations

Lim. nr	Limitation description
1	Users have to cope with a very large number of different identities and associated credentials.
2	Users must handle several different types of authentication systems.
3	It is a expensive model since it works as a single non shared service.
4	Lack of standards, resulting in security inconsistencies and lack of interoperability between systems.
5	Lack of interoperability between different systems.
6	Users do not have control over their own privacy.
7	No Federation. Each Identity provider has its own scheme.
8	Relies on physical identity elements.
9	Easy to compromise since an attacker can more easily associate digital personas to real life identities.

2.1.3 Federated Model

An Identity federation can be defined as the set of agreements, standards and technologies that allows a group of distinct administration domains to negotiate the establishment of a circle of trust [36] where they can then mutually recognise the user identifiers that are managed and entitlements that are issued by the other federation members they trust. Within a federation, several identity providers (IdP) can be used to provide a valid identity to any service domain that falls within the federation and authentication within the federation is achieved by presenting a valid identifier emitted and authenticated by an IdP that is trusted by the federation. However each service domain is still free to differentiate the level of service provided to his own managed identities from the level of service and resources that are provided to other identities managed by the other distinct federated administrative domains.

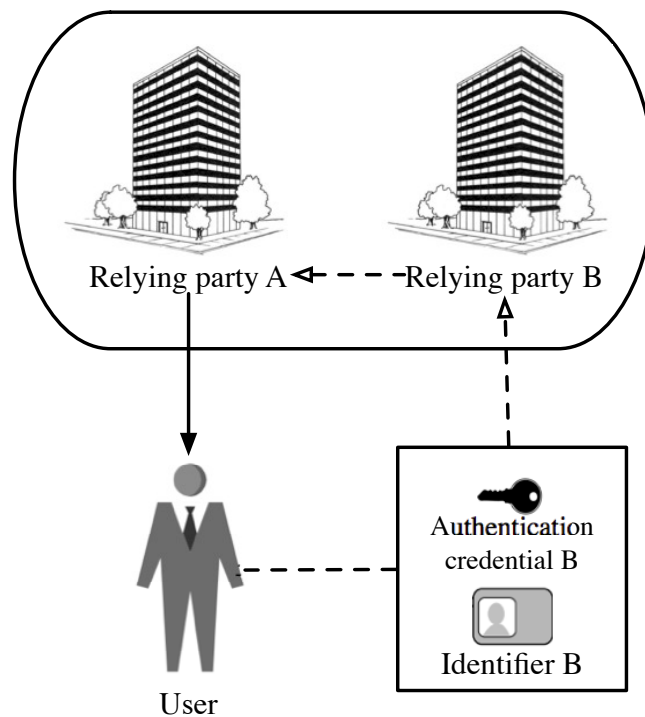


Figure 2.1: Federated model example

In a federated identity domain, agreements are established between the participating members to determine which identity attributes need to be recognised across all the different participating federated domains. These agreements include well defined policies and technology standards that allow for the mutual interoperability of the participating federated members. A connection can also be established between the

different identifiers owned by the same user in different domains thus linking these identities within the federation. This results in a single virtual identity domain where a user authenticated by a single identity provider using one of his identifiers, can be considered to have been identified and authenticated within all the other identity providers within the federation as well. This can be technically achieved by passing assertions between service providers. (see figure 2.1)

Sending an assertion does not require the users' credentials. The acceptance of user access assertions from one identity provider to another is based on the trust established by the adherence to the circle of trust common policies. A federation of isolated identifier domains gives the client the idea that there is in fact a single identifier domain. The user can still hold separate identifiers for each identity provider, however, he does not necessarily need to know and manage them all. A single identifier and credential is sufficient for him to access all services within the federated domain and this can be used to provide a Single-Sign-On solution for the entire federation.

2.1.3.1 Shibboleth

Shibboleth [37] is a very popular, widely deployed and standardized single sign-on federated platform for the web. It has been developed in the United Kingdom by the Joint Information Systems Committee (JISC) to support federated identity management across all educational institutions. Shibboleth deploys a wide variety of standards, like for example the OASIS Security Assertion Markup Language (SAML) for attribute assertions, in order to provide a federated single sign-on environment and a secure attribute exchange mechanism for web applications.

In shibboleth, access control is based on a set of identity attributes provided by a trusted identity provider and a set of rules defined by the relying party the user wants to interact with. If the users' attributes positively match the relying party's rules, access is granted. The basic premise lies on the existence of a circle of trust, where the federated institutions trust the assertions provided by every member of the federation.

Shibboleth main objective is to have everyone involved in the educational system (elementary school to higher education) enrolled into a wide circle of trust, in order to better promote the sharing of online educational resources.

2.1.4 Centralized Model

In centralised models, there exists a single identity manager that is used by all relying parties, either exclusively, or in addition to other security domains identity managers.

The concept of single sign-on [6] was introduced by this model in order to create an integration solution for the still highly popular traditional based applications. There is only one centralized identity provider that is trusted by all the participating relying parties to manage the user credentials necessary to gain access to their resources. The user only needs to authenticate once to gain access to the services provided by the trusting relying parties. (see figure 2.2)

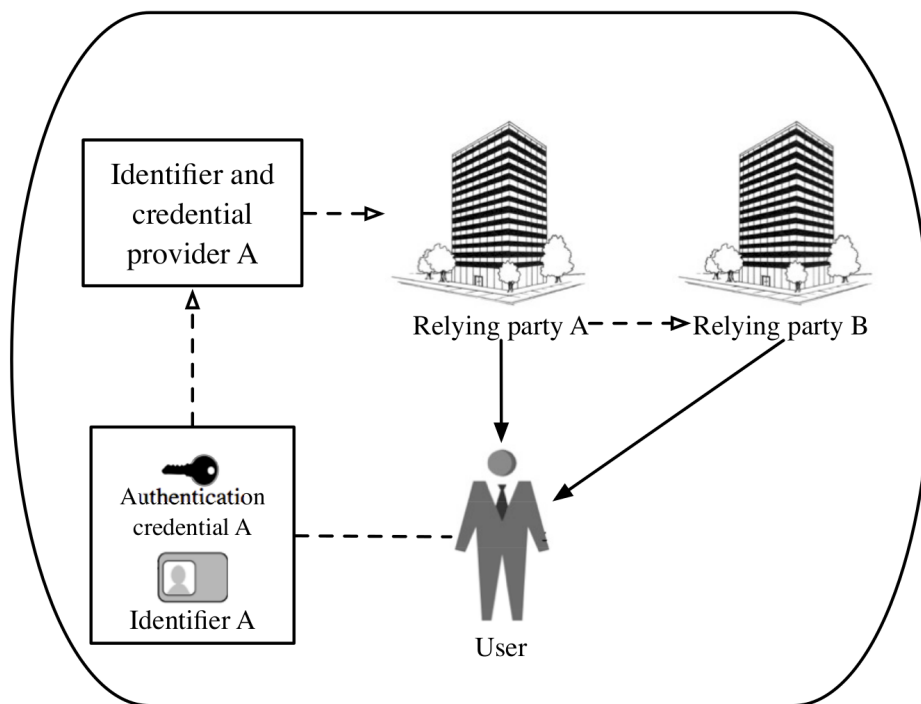


Figure 2.2: Single sign-on model example

This Single Sign-On scenario is very similar to the previously described federated identifier scenario, except that no mapping of user identifiers is needed because the same identifier is being used by every relying party. For example Kerberos based authentication solutions, where the Kerberos Authentication Server acts as the centralised identity provider, belongs to this category (Kerberos was described in subsection 2.3.1). We now proceed to describe other identity management solutions that may fall within this category.

2.1.4.1 OpenID

OpenID is an open standard for identity management and user authentication that was originally conceived to be deployed on an highly decentralized manner. The idea was to provide each user with the possibility of deploying his own OpenID identity provider and thus establish a truly user centric, massively distributed, self assembly identity infrastructure for the Internet. However it did not work this way. Big Internet identity providers like for example *Google* and *Facebook* started to employ OpenID as the SSO mechanism to uniform access to their own services. This resulted in the construction of massive OpenID based identity management systems with hundreds of millions of users that they then opened up for other relying parties, belonging to other administrative domains, to use. In practice this ended up transforming OpenID, the paradigm of decentralization, into a protocol that is nowadays mainly employed in the Internet to make use of the Identity services provided by Google or Facebook, in a truly centralized way.

The OpenID protocol [2] allows a user to sign-in into very distinct domains with the same OpenID identifier (a URL) and at the same time control what subset of his identity attributes will be disclosed to each one of the relying parties he logs in to.

In order for a user to be authenticated with OpenID at a relying party, he starts by having his browser redirected to his OpenID provider by the relying party, where he is then authenticated (usually by login/password). If the authentication process is successful the user is then asked to authorise the OpenID provider to disclose the identity attributes that are being requested by the relying party, after which the user is once again redirected to the originating requesting relying party where access is finally granted.

Unfortunately the full set of standardised and universally recognised identity attributes for OpenID is unfortunately very small. This decreases the usefulness of the protocol and has so far limited its deployment almost exclusively to the simple authentication domain. More recently the OpenID foundation has started to work on a new protocol named OpenID Connect [18] that aims to unify authentication and authorization into a single service protocol. This unification creates the right means for data access authorization and it is a firm step towards solving the issues resulting from a too limited set of widely recognized identity attributes.

2.1.4.2 Microsoft .NET Passport

.NET Passport is a single sign-on identity management system developed for e-commerce by Microsoft, where email addresses have been adopted as the user main identifier. In .NET Passport, credential issuance and authentication are fully centralised operations under the strict control of the Microsofts identity provider.

Microsoft extended the centralized model with the creation of "InfoCards" [38] (later named *Windows CardSpace*) in order to ensure a better level of security and trust between the involved parties however in 2011 the Windows CardSpace was discontinued. Microsoft is currently working on a replacement called U-Prove [39].

U-Prove aims to provide a multiparty security (issuing organizations, users, and relying parties can protect themselves not just against outsider attacks but also against attacks originating from each other) and at the same time enables customized privacy settings (including authenticated anonymity and pseudonymity).[40]

2.1.5 User Centric Models

User centric identity management is currently a hot research topic. Its main objective is to empower users by returning control about who can access their personal data back to them, the truly legitimate data owners [41, 12]. User centric identity management models provide the users with the means to effectively decide who can access their identity data. The whole process is based on having an identity manager receiving requests for user attributes made by the data requesters, the relying parties, for which the user is then provided with the means to choose the most appropriate subset of identity attributes to disclose. This gives end users more privacy control and at the same time more responsibility over their own choices.

For example Figure 2.3 illustrates a user centric scenario where users can store and manage their identifiers and credentials, from different service providers, with a smart-phone. The user decides which identifier and identity attributes will be deployed to the a service provider and then inserts a pin to authorize the operation.

2.1.5.1 Higgins

Higgins was initially known as the Eclipse Trust Framework[42]. Its main objective is to establish a software layer that can be built upon and expanded with other

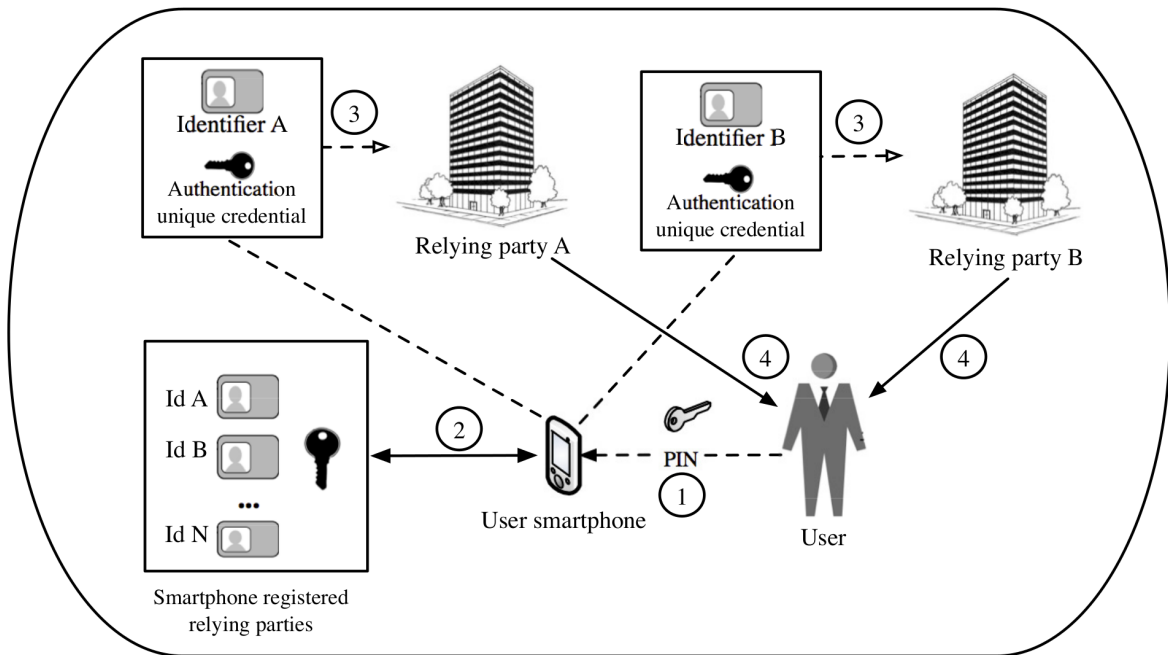


Figure 2.3: Smartphone user centric model

components and adapters. The Higgins Project splits identity information into small blocks of data, known as i-cards, which can then be directly controlled by the user and are meant to replace traditional password-based login systems.

I-cards can be created by mixing different digital identity data from different domains. By mixing these different contexts, data aggregation is accomplished and interoperability between distinct domains is thus established. The Bridging contexts is done by a user agent operated by a browser extension and the Higgins card selector service. Since the users' data is directly processed at the user agent, the browser, identity information does not need to be shared with a trusted third party and interoperability can be reached by just having the user fully managing his own data.

2.2 Attribute Aggregation

Attribute aggregation is the process of acquiring identity attributes from multiple distinct identity providers and Attribute Authorities (AA) in a single session. An AA is an entity responsible for the security, management and disclosure of users data.

The users attribute aggregation deployment scenarios can be quite involved and com-

plex and we were therefore required to first delve into a more comprehensive exploratory explanation of the different ways this has been described in the literature to better understand the existing models and to start providing some answers to some interesting but difficult research problems like: How much personal information the relying parties/service providers (RP/SP) and Identity Providers should have access to? Who should be responsible for the final aggregation of the users' data? What authorizations and proofs are necessary to securely request a set of personal data? How can a RP/SP be sure that who provided the data is really its legitimate Attribute Authority? How can the RP/SP know that a complete set of attributes relates only to one single identity?

After a comprehensive research about what are the most popular and comprehensive user-centric attribute aggregation models, we came to the conclusion that the most relevant models are:

2.2.1 Client mediated assertion

Based on an intelligent browser user agent that guide the user to the different identity providers, obliging the users a high level of interaction by authenticate themselves on each identity providers. The browser user agent is responsible for the attribute aggregation and the attribute delivery to the requester relying party. In this model no relying party or third trust party is involved in order to request and aggregate the information its only based on browser technologies. [43]

2.2.2 Identity Federation model

Require a federated network in order to operate, during users' authentication process, a secret is generated and shared between all federated identity providers by a user agent. The first contacted identity provider provides to the requester relying party the details of the others identity providers that falls into the federation thus establishing the necessary means to the relying party request the necessary attributes. [44]

2.2.3 Relying Party mediated attribute aggregation

Its based on the identity federated model however discards the necessity of a federated network. The relying party redirects the user-agent to each identity provider thus

obliging the users to a high level of interaction since the user has to authenticate himself in every identity provider. The browser user agent is the responsible for attribute aggregation thus requiring specific browser based technologies. [45]

2.2.4 Identity proxying/chaining

The relying party has to fully trust in a single master federated identity provider that is responsible to request and aggregate all requested attributes before send it back to the requester relying party. In other words the master identity provider can request attributes from all others identity providers that are within its federation circle.

This model have a low level of protection for user attributes since every intermediary identity provider must relay on credentials intended for a third party. This means create a possibility for a substitution attacks and as well as the possibility of an identity provider sending false authentication information. [46]

2.2.5 Identity Relay

The relying party trust in a single master federated identity provider that is responsible to request all attributes to the relying party, these attributes are returned directly to relying party, in other words relying party is responsible to aggregate the set of necessary attributes.

This model is like identity proxying but with a reduced level of trust on master identity provider since the attribute aggregation is done by the relying party. The master identity provider only redirect the attribute requests the others identity providers that answer directly to the relying party. [43]

2.2.6 Linking Service

In this model only the user knows about all his identity providers. A service called linking service is responsible to hold minimal information that allow relying parties to obtain their queries from different identity providers domains. After a user authenticates, the identity provider offers the possibility of attribute aggregation and if the user authorizes it, the information to access the linking service is shared with the relying party. The aggregation of attributes can be done by the linking service itself

Table 2.3: Information Security Attributes

Key aspect	Description
Confidentiality	Information is shared only among authorised persons or organisations.
Integrity	Information is authentic, complete and can only be accessed or modified by those authorized to do so.
Availability	Assurance that the systems responsible for delivering, storing and processing information are accessible when needed.

or by the relying party.

The linking service model offers a good level of privacy for user attributes, since all attribute assertions are signed by their sources, and users permission are requested in order to deploy the attributes. The relying party must have a high trust on the linking service node. [47]

2.3 Identity Management Common Standards

In order to establish the right means to guarantee key aspects of information security (see table 2.3) in identity management systems, several standards started to appear, some of them nowadays became essential tools in the deployment of identity management systems.

In this section we described the most common standards used in identity management systems and give a view of a new and still on development named OIX.

2.3.1 Kerberos

Kerberos is one of the earlier security standards developed. Kerberos is a secure method for authenticating a request for a service in a computer network. Kerberos was developed in the Athena Project at the Massachusetts Institute of Technology (MIT). [48]

Kerberos allows a user to request an encrypted "ticket" from an authentication process that can then be used to request a particular service from a server. In Kerberos the users' password is computed by an one-way hash function that is used as a symmetric key to read the "ticket" provide by the kerberos server in order to establish the

authentication. Notice that the users' plaintext password do not pass through the network.

2.3.2 Security Assertion Markup Language (SAML)

Security Assertion Markup Language is a XML-based open standard data format for exchanging authentication and authorization data between an identity provider and relying parties, in other words, SAML defines an XML-based secure framework for exchanging identity information across security domains for purposes of authentication, authorization and single sign-on. [49]

SAML specifies four components (see image 2.4): (1) The assertion component where it is divided into three types: (a) the authentication assertion to validate the users' identity, (b) the attribute assertion to hold specific attribute about the user, (c) and the authorization assertion to identify what the user is allowed to do; (2) the protocol component to defines how SAML requests for and receives assertions; (3) the binding component to defines how SAML message exchanges are mapped into standard messaging and communication protocols like HTTP, SMTP and also SOAP; (4) and the profile component describes in detail how SAML assertions, protocols, and bindings combine to support a defined use case, the most well known SAML profile is the web browser single sign-on profile.

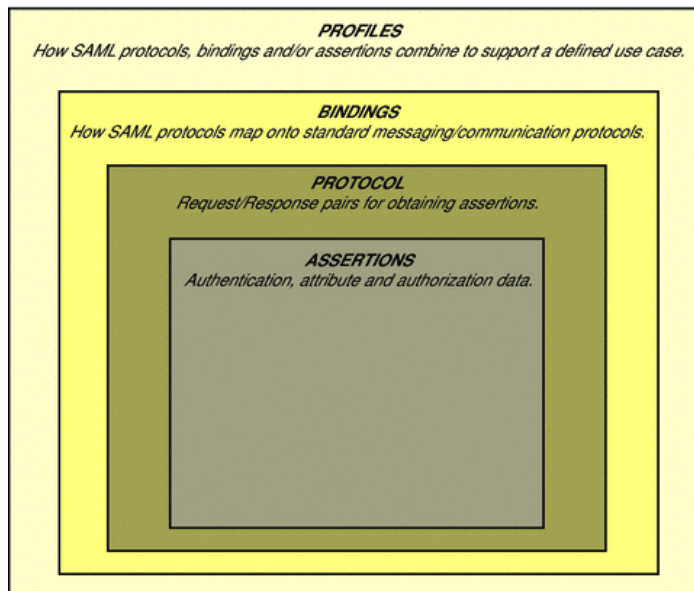


Figure 2.4: Components of the SAML Specifications

The SAML specification defines three roles: the principal (usually the user), the

identity provider, and the service provider/relying party. In the use case addressed by SAML, the user requests a service from the service provider that requests and obtains an identity assertion from the identity provider. On the basis of this assertion, the service provider can decide whether to perform some service for the connected user. Before delivering the identity assertion to the service provider, the identity provider may request some identification or authentication from the user such as user/password login scheme in order to authenticate the user.

2.3.3 Extensible Access Control Markup Language (XACML)

XACML is an initiative to develop a standard for access control and authorization systems by defining a declarative access control policy language implemented in XML and a processing model describing how to evaluate authorization requests according to the rules defined in policies. In other words XACML not only processes the authorization requests, but it defines the mechanism for creating the complete infrastructure of rules, policies, and policy sets to arrive at the authorization decisions. [50]

In XACML the access requests, are sent to a Policy Enforcement Point (PEP), located at a web server, which creates a XACML request and sends it to the Policy Decision Point (PDP). The PDP determines the answer based on the existing policies in the policy repository and sends back its determination to the PEP. The response can be either access permitted or denied, with the appropriate obligations. Obligation is a directive that if applicable must always be carried on the authorization process.

2.3.4 Web Services Federation (WS-Federation)

WS-Federation [51] describes how the assertions transformation model inherent in security token exchanges can enable trust relationships and federation of services. This allows scenarios where authorized access to resources managed in one realm can be provided to security principals whose identities and attributes are managed in different realms.

Web Services Federation has mechanisms for identity brokerage, attribute discovery and retrieval, authentication and authorization assertions between federation members, and protecting the privacy of these assertions across federation boundaries.

These mechanisms provide a basic trust model between identity providers and de-

pendent parties through WS-Security, WS-Trust, and WS-Security Policy: (1) WS-Security is a communications protocol and is the component that allows secure access from web clients using HTTP or from web services clients directly. (2) WS-Trust is the component that defines the Security Token Service (STS) and the protocol used to request or issue a token. These tokens are generated from the assertions about an entity requesting a service, (3) and WS-SecurityPolicy describes the STS policies and its associated assertions.

2.3.5 Digital ID wallet

Digital ID wallet is an ITU-T international standard [52] that is a web-based digital wallet. It is employed as storage for private data such as addresses, users' ID, and passwords. This standard follows the user-centric method diffusing the user empowerment concept facing the federated identity management systems that move the users' control to the service providers and identity providers. The Digital ID wallet standard emerges to shift back the control to its right owner, the user.

This standard uses tokens as secret keys that are automatically created for each visited web site, which is stored in a client-based digital wallet. Thus, it is not necessary for users to memorize individual passwords. The Digital ID wallet enables the user to register and log onto a website, store personal information, and other data at any time. This approach is independent of the existence of a Certification Authority and is well suited to support mobile communication.

2.3.6 Open Identity Exchange (OIX)

The Open Identity Exchange [53] is a non-profit corporation serving as an independent, neutral provider of certification trust frameworks for open identity technologies. This means it follows an open market model to provide the certification services requested to guarantee the necessary levels of identity assurance and protection. In other words OIX is an open identity trust framework provider that provides certification in order to enables a relying party to trust the identity, security, and privacy assurances from an identity provider.

The OIX objective is to reduce the friction of using the Web by solving problems like: how does a relying party know it can trust credentials from an identity service provider without knowing if that providers' security, privacy, and operational policies

are enough to protect the relying party's interests? OIX states that problem is a business, legal, and social problem and not a technological problem.

Despite the fact that OIX is still on a modelling stage this project looks really promissory since companies like Google, PayPal and at&t quickly become associated.

2.4 Security Concepts

In order to establish a secure connection to exchange information guaranteeing both users are who they say they are and the messages are not being attacked/manipulated secure methods/protocol were used. In this section some relevant technologies/protocols are revised.

2.4.1 Public Key Cryptography

Public key cryptography is an asymmetric scheme that uses a pair of keys for encryption and signature process. The public key is employed to encrypt data or verify digital signatures, and the corresponding private key is used to decryption or sign data. (see figures 2.5 and 2.6) The public key must be disseminate over the global key servers while the private key must be highly secured stored by the user. Anyone with a copy of a public key can then encrypt information that only the respective private key owner can read.

It is computationally unfeasible to deduce the private key from the public key so if the private key is lost its public key is lost too. The public key cryptography solved the need for the sender and the receiver to share secret keys via some secure channel, allowing users to exchange messages securely without the need of a pre-existing security arrangement. The most well known public key cryptography systems are:RSA, Diffie-Hellman and DSA. [54]

2.4.1.1 Public Key Infrastructure

A Public Key Infrastructure contains the certificate storage facilities of a certificate server, but also provides certificate management (the ability to issue, revoke, store, retrieve, and trust certificates). The main feature of the Public Key Infrastructure is the introduction of what is known as a Certification Authority (CA), a person, group,

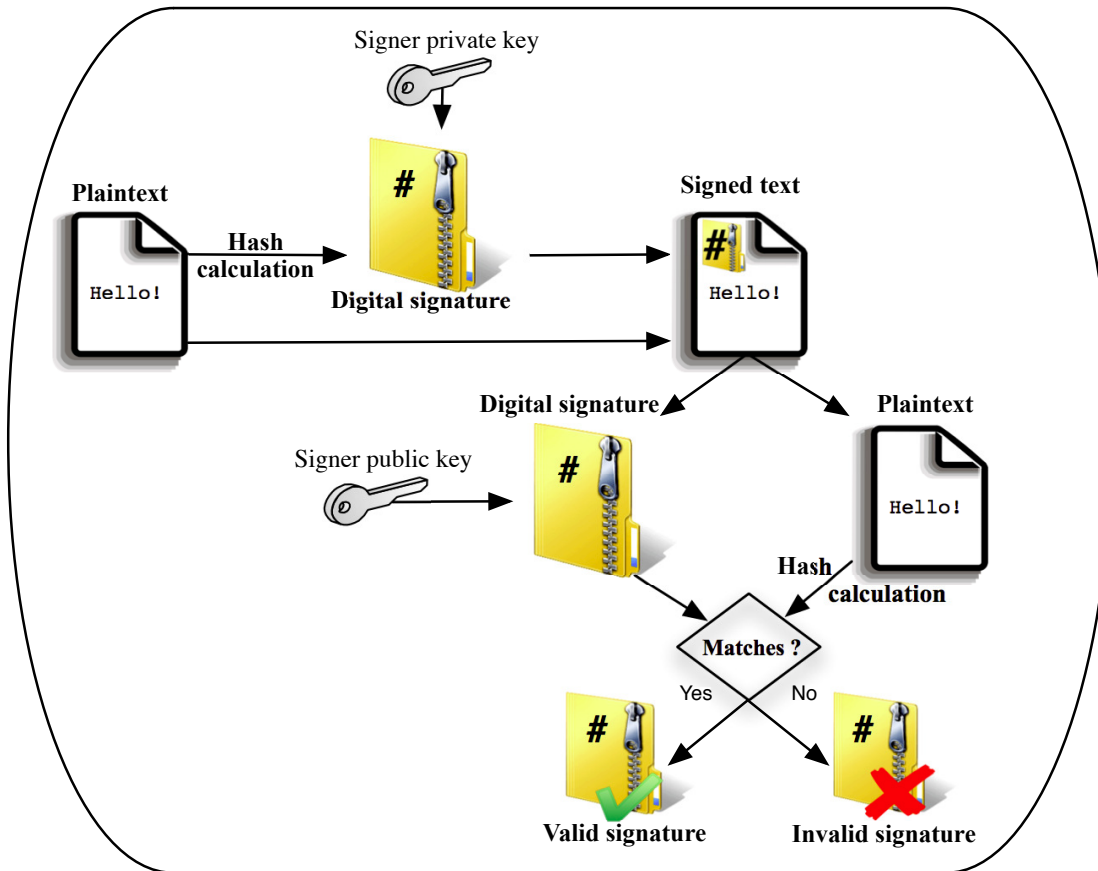


Figure 2.5: Public key cryptographic signature process

department, company, or other association that an organization has authorized to issue certificates to its computer users. The certification authorities are analogous to a country's government's Passport Office and can be seen as a trusted third party that creates certificates and digitally signs them using its private key. Because of its role in creating certificates, the CA is the central component of a Public Key Infrastructure. Using the CA's public key, anyone wanting to verify a certificate's authenticity verifies the issuing CA's digital signature, and hence, the integrity of the contents of the certificate. (see figure 2.7)

2.4.1.2 Digital Certificates

Digital certificates, or certs, simplify the task of establishing whether a public key truly belongs to the purported owner.

A certificate is a special form of credential. Examples might be your national iden-

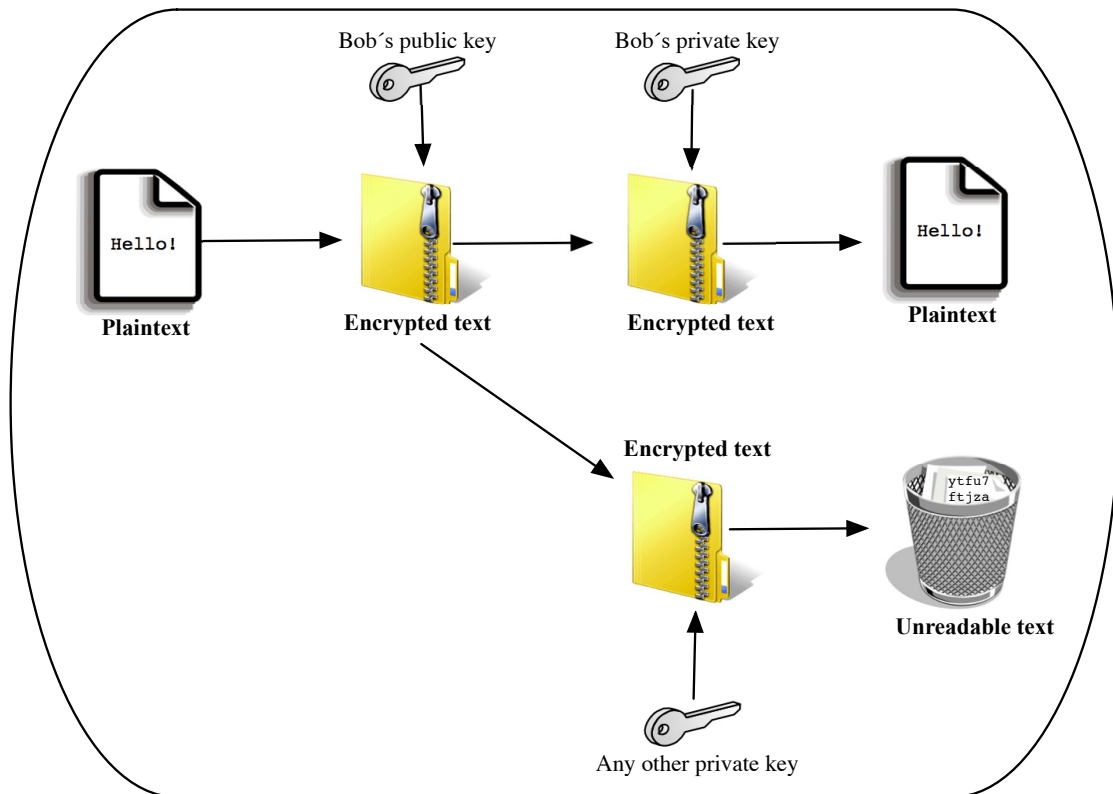


Figure 2.6: Public key cryptographic encryption process

tification card, your drivers license, or your birth certificate. Each of these has some information on it identifying you and some authorization stating that someone else has confirmed your identity. Some certificates, such as your passport, are important enough confirmation of your identity that you would not want to lose them, lest someone use them to impersonate you.

A digital certificate is data that functions much like a physical certificate. A digital certificate is information included with a person's public key that helps others verify that a key is genuine or valid. Digital certificates are used to thwart attempts to substitute one person's key for another.

A digital certificate consists of three different components:

1. A public key.
2. Certificate information. (Identity information about the user)
3. One or more digital signatures. (Depending on the certification type)

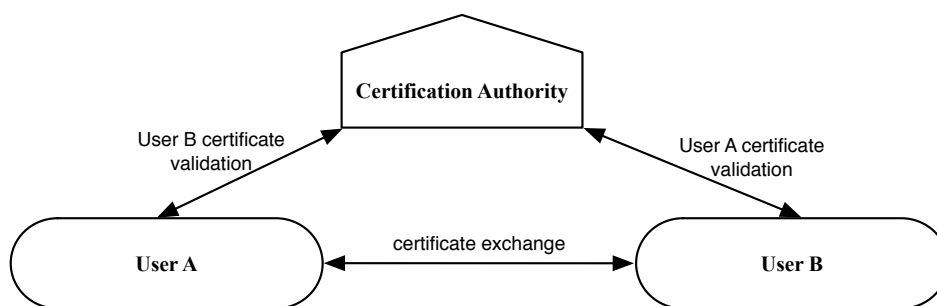


Figure 2.7: Certification validation by the usage of a Certification Authority

The purpose of the digital signature on a certificate is to state that the certificate information has been attested to by some other person or entity. The digital signature does not attest to the authenticity of the certificate as a whole; it vouches only that the signed identity information goes along with, or is bound to, the public key. Thus, a certificate is basically a public key with one or two forms of ID attached, plus a stamp of approval from some other trusted individual.

Certificates are only useful while they are valid. It is unsafe to simply assume that a certificate is valid forever. In most organizations and in all Public Key Infrastructures, certificates have a restricted lifetime. This constrains the period in which a system is vulnerable should a certificate compromise occur. There are also situations where it is necessary to invalidate a certificate prior to its expiration date, such as when the certificate holder terminates employment with the company or suspects that the certificate's corresponding private key has been compromised. This is called revocation. The most well known certificates are the PGP (pretty good privacy) certificate usually used email systems and the X.509 certificate massively deployed on the web in order to establish the HTTPS protocol.

The most well recognized PGP certificate characteristics are:

- Is issued by its creator (self-signed)
- Support multiple signatures in order to grant a greater trust on network
- Based on a web of trust
- It is free

and the most well recognized X.509 certificate characteristics are:

- Is issued by a certificate authority (CA)
- Hierarchical approach
- Supports only one signature usually from the CA
- In order to be well recognized requires some costs.

2.4.2 Smartcard

A smartcard is a pocket-sized device with an embedded microprocessor that can provide secure: identification, authentication, data storage and application processing. The chip of the microprocessor guarantees tamper-resistance [55] and its protocol interface assures the security over its data access by being logically impossible to extract information without the appropriate keys. The protocol interface set a strict control over what can be directly accessed from the smartcard (even with the appropriate pin) making almost impossible to clone it.

2.4.2.1 Mobile Secure Card

Due to its potential economical factor [8], the hunger for mobile devices that can act as an authentication/authorization node are daily increasing. Mobile operators started to explore the usage of smartphone as authorization brokers. Despite the fact that mobile operator have the best profile to provide a service like that since they already have a whole system prepared for this means, this operations will require an extra fee for their customers. In order to not rely on single mobile operators and flee from the extras fees an alternative path is the usage of smartcards on the smartphones.

Nowadays almost all smartphones accept the microSD card as its standard memory extension to expand its storage capacity. In order to escape from the mobile operator taxes we relied on the security properties of the mobile security card [56] for mobile devices. This card provides an interesting technical standard known as SmartSD which provides the necessary crypto components and device physical non-tampering for our architecture. This process is archived by adding a smartcard component besides the flash component inside the SD card. The mobile security card is a microSD card that explores the SmartSD standard by embedding a smartcard chip that uses JavaCard OS. This card is responsible for guarantee a strong users authentication and trustworthy protection of data by the usage of cryptography components.

2.4.3 Valet key based protocol

Nowadays many common authorization protocols are based on a valet key concept. They all employ a token as a secure digital object that a pre-authorized entity needs to present in order to have direct access to some restricted resource. In other words, these tokens look like a key for data access in the sense that any entity who possesses the key has temporary and restricted access to the protected resource.

The most common scenario is a token based authorization scheme involving three distinct actors: The data owner (User), a third party application (Relying Party) and the user data storage (Attribute Authority). In this scenario a user wants to provide a relying party with an authorisation to access his data that resides on a certain attribute authority. To achieve this, the relying party redirects the user to the attribute authority with a formalised request where the user is asked to authorise it, this request includes the data that the relying party desires to obtain and for how long time he wants to access it. After authorisation, the attribute authority returns to the relying party a signed authorisation token that allow the relying party to access the requested data by presenting the signed authorisation token while it remains valid. These tokens can be revoked at any time by the user that owns or manages the data.

These tokens must be extremely difficult to falsify and provide a flexible security mechanism for the attribute authorities to more easily manage access control to restricted resources. At the same time these tokens provide the Relying Parties with the means to access otherwise restricted resources without the need to obtain, share and manage other types of credentials like login/passwords.

2.4.3.1 Open Authorization (OAuth)

OAuth is an authentication and authorization protocol originally developed for web applications that provides a standard method for clients to access server resources on behalf of a resource owner by the usage of tokens. It also provides a process for end-users to authorise third-party access to their server resources without sharing their credentials, using user-agent redirection [3]. The most common analogy to this protocol is the valet key.

2.4.3.2 Azure Microsoft

Windows Azure is a Microsoft's application platform that implements a public cloud. Windows Azure can be used to build a web application that runs and stores its data in Microsoft datacenters. It can connect on-premises applications with each other or to map between different sets of identity information. In order to allow federation with different identity providers such as a corporate Active Directory, Windows Live ID, Facebook, Google, and OpenID 2.0 identity providers. Microsoft created an access control service named as Windows Azure AD Access Control that work as a cloud service that provides the necessary Security Token Service (STS) in order to establish the network federation between the different identity providers.

2.5 Communication Concepts

In order to establish a connection between two or more different nodes a communication channel must be established. This section presents different protocols used for the communication channel creation. This section describes a messaging protocol named as XMPP, the well known HTTP by the usage of web services and an innovative analogue communication channel archived by the usage of QR-Codes.

2.5.1 XMPP

XMPP is an open technology for real-time communication that uses the eXtensible Markup Language (XML) as a base format for exchanging information encapsulated into pieces of XML documents. These XML documents are sent from one entity to another [57] by using an appropriate application level transport protocol according to the network availability. XMPP servers provide a numerous set of standard of services that can be adopted by the most different types of applications.

The XMPP is the almost ideal communication infrastructure for the establishment of a security channel because of its services, namely:

- Almost real time messaging, essential for critical services.
- Authentication by the usage of certificates, guaranteeing a high level of trust and non-repudium.

- its ability to efficiently operate over HTTP by the means of the BOSH (Bidirectional-streams Over Synchronous HTTP) protocol [58], where two non directly addressable devices located on private closed intranets with minimal Internet access, can locate each other over the Internet and then freely exchange messages between themselves in a reliable and safe way.
- Its capacity to store and forward messages in case any of the nodes becomes offline, which is proving to be essential for asynchronous communications.
- Its scalability to avoid bottleneck problems and the fact that it is a mature fully supported and approved Internet standard, widely deployed and an important part of the communication operations and infrastructure of large distinct companies like: Google, Facebook, Blizzard and Steam.
- The possibility of multi sessions in a single XMPP account organized by the creation of XMPP resources.
- Peer-to-peer media sessions: allows a peer to negotiate and manage a composed of large and complex audio/video data streams with another peer.

2.5.2 Web service

The term web service have many imprecise and ambiguous definitions. This issue is derivable from the different existing concepts of the web services. In general a web service can be defined as a distributed software system designed to support interoperable machine-to-machine interaction over a network that usually interacts with the World Wide Web (WWW) typically delivered over Hyper Text Transport Protocol (HTTP). The communication payloads of web services are usually XML documents, however, when performance demands web services can also use binary payloads. The web services can be identified in two major classes:

2.5.2.1 REpresentation State Transfer (REST)

REST is a style (REST-style, also known as RESTful web service) of software architecture for networks connected through hyperlinks. The most well known applied domain for REST-style is the WWW that is a distributed network of hypermedia. In order words, the RESTful web service is a subset of the URI (Uniform Resource Identifier) in which provides an uniform interface semantics to manipulate web resources. [59]

The main aspects of REST are:

- Resource Identification through the usage of URI
- Uniform Interface for all resources (HTTP as the Application-level Protocol, observe table 2.4):
- Hyperlinks to define relationships between resources and valid state transitions of the service interaction.

Table 2.4: HTTP verbs and their meaning in RESTful web servers

HTTP verb	Meaning in RESTful
GET	Query a resource in web server.
POST	Create a resource in a web server.
PUT	Update an existing resource in a web server.
DELETE	delete a resource in a web server.

2.5.2.2 Simple Object Access Protocol (SOAP)

SOAP is a lightweight XML-based messaging protocol that is independent of any platform, transport protocol or operating system. It defines a set of rules for structuring XML messages that can be used for simple one-way messaging. This set of rules are known as SOAP message. It is not tied to any particular transport protocol though HTTP is the most popular. Nor is it tied to any particular operating system or programming language so the clients and servers in these dialogues can be running on any platform and written in any language as long as they can formulate and understand the SOAP messages. In SOAP-based web services the information remain "outside of the web" since there is no specific URI for the web server resources. [59]

There are two types of SOAP requests:

- **SOAP-RPC** that is an implementation of a Remote Procedure Call (RPC). The clients invoke the web service by sending parameters and receiving return values. RPC-style web services follow call/response semantics; therefore, they are usually synchronous, which means that the client sends the request and waits for the response until the request is processed completely.

- **Document style message**, the client and/or server passes an XML document as the body of the SOAP message instead of parameters (like SOAP-RPC). Since the document message is usually a self-contained XML, it is better suited for asynchronous processing considering that commonly there is no wait for a response.

2.5.3 Quick Response code

The Quick Response code (QR code) is a two-dimensional square shape that encodes a reasonable amount of digital information into a small amount of 2D space. The encoding is achieved with the careful positioning of varying size black and white smaller squares within the 2D space defined by the QR square.

These 2D codes are normally displayed within web pages or printed in paper posters and are employed to quickly exchange digital information with mobile devices that would otherwise had to be entered by hand. This is accomplished by having the mobile device to digitally scan and decode the displayed QR code with its built-in optical camera [60].

The usage of QR codes to share secret information can, in a way, be seen as the establishment of a rather new secure communication channel that takes advantage of the analog security properties of the optical channel that is employed during the scanning of the QR codes by the smartphone. In practice QR codes are used to simplify and make practical the enrollment process between different applications.

Chapter 3

Architecture Components

In this section we describe in detail the main technological components we have employed in OFELIA. We also discuss the main aspects behind some of the alternatives and compromises we had to make to integrate our vision with already existing real world services and devices (ex: Google XMPP infrastructure, Android devices, etc...). We also take some time to describe the conceptual data model for attribute aggregation and its most relevant aspects like the protocols and services we have employed to integrate the different components that compose the proposed architecture. Figure 3.1 shows the main relationships between the principal components and the type of communications and the data exchanges that can occur between them in a simplified way.

3.1 Architecture Technologies Overview

In what follows we provide a more detailed description of the functional role played by each of the main technology in our proposed architecture.

3.1.1 microSD mobile card

The mobile security card is a microSD card that explores the SmartSD standard by embedding a smartcard chip with JavaCard OS. This card have a special place in our architecture since its responsible for guarantee a strong users authentication and trustworthy protection of data. Otherwise we would to rely on a regular file based

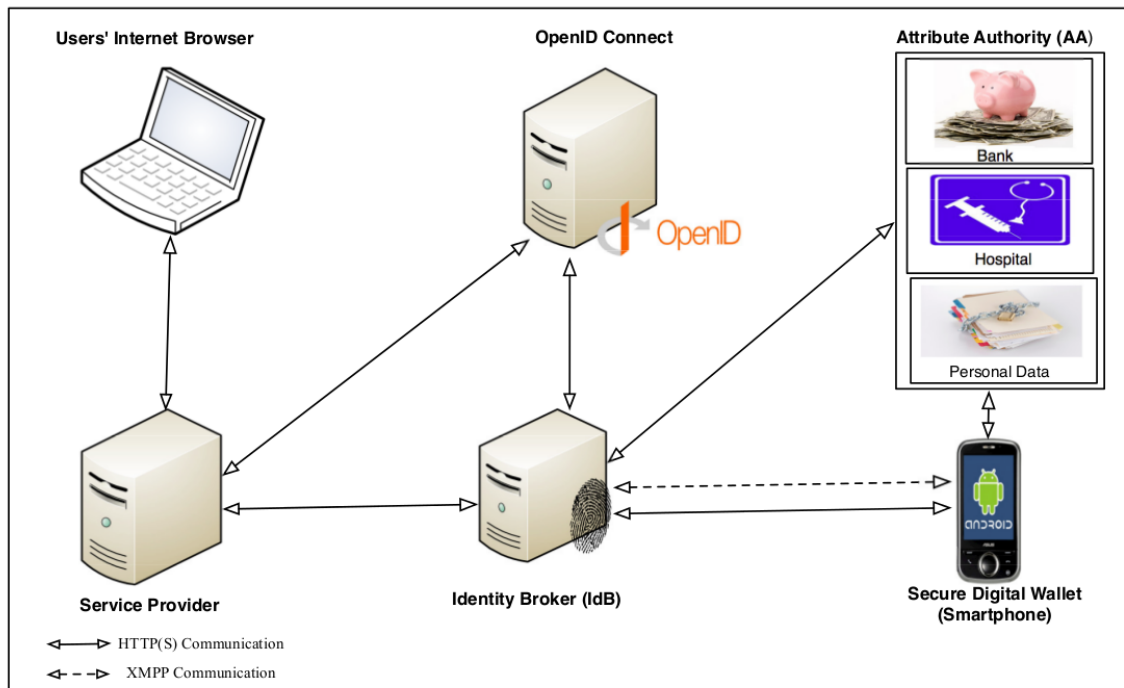


Figure 3.1: The proposed architecture communication

keystore, turning the smartphone in a desirable target of attacks where the keystore file would be easily compromised. So it is reasonable to put the file based keystore level of security in tandem with the security provided by a much simpler login/password based scheme. In fact an attack on a password protected keystore involves a password guessing attack completely analogous in terms of complexity to what happens with an attack directed towards a login/password scheme, the only thing really different in this case being the need to possess a copy of the keystore file in order to proceed with the attack.

3.1.2 The XMPP messaging protocol

Arguably, in the mobile world, there is some difficulty in directly addressing and communicating with Internet enabled mobile devices. In the mobile world an implicit direct communication with the device is almost impossible due to the shortage of public IPs addresses faced by Internet service providers and mobile operators. In the future, IPv6 is supposed to solve this problem, however it is our strong belief that the mobile Telecommunications operators will still not allow this kind of direct communication to mobile phones due to their very inflexible business plans, where the mobile phone

is nowadays mostly regarded simply as a consumer device and never as a provider of services. In fact Telecommunications operators restrict even the ports available to initiate communications and the most restrictive only allow direct communication with the Internet over port 80 (standard HTTP port).

A neutral rendezvous point on the Internet where our architecture nodes can meet to exchange messages is thus obviously necessary. Towards this end, XMPP messaging is proving to be an almost ideal communication infrastructure for OFELIA to circumvent these communication restrictions because of its ability to efficiently operate over HTTP by the means of the BOSH(Bidirectional-streams Over Synchronous HTTP)[58] protocol where two non directly addressable devices, located on private closed intranets and with minimal Internet access, can locate each other over the Internet and then freely exchange messages between themselves in a reliable and safe way.

3.1.3 Secure access authorization tokens

For security reasons the secure access authorisation token must be very hard to falsify. In our project it takes the form of a base64 encoded XML excerpt, containing a hash value created from elements for a large pseudo random number [61] and a simple semantic statement element, describing the authorization validity restrictions that apply to this particular authorization. This statement can express for example temporal restrictions. In order to ensure a right level of authentication and non-repudiation, this XML excerpt is always digitally signed by the users' smartphone private key. The resulting XML document is then encoded into a base64 string, which then constitutes a well formed secure access authorization token. These tokens provide a very flexible security mechanism for the attribute authority to more easily manage access control to restricted resources. At the same time these tokens provide the relying parties with the means to access otherwise restricted resources without the need to obtain, share and manage other types of credentials.

In our proposed architecture these authorisation tokens are issued by the authority broker (users smartphone) and are shared with the Identity Broker and the Attribute Authority, in order to provide for data access. It is also important to clarify that in our model the user maintains the revocation rights by being able to unconditionally revoke these tokens, at any given moment, by the means of his personal smartphone that acts as an Authority Broker.

3.1.4 The OFELIA TRUST infrastructure

One of the key critical components of our proposed architecture is the management of trust among the participating components. To establish the necessary level of trust we rely on a Public Key Infrastructure (PKI) that is responsible for the management of the certificates that are at the core of the privacy, trust, non-repudiation and authentication infrastructure mechanisms that we need to put in place to secure our architecture.

To establish a stronger and therefore more trustworthy identity/authentication between the different actors, namely: the relying party (data requester), the attribute authority (data storage), the identity broker (identity manager) and the authorization broker (users smartphone), we rely on the deployment of a well managed standard compliant PKI that can also sign PGP (Pretty Good Privacy) and X509 certificates. These certificates are then used as securely vouched identity credentials that is employed to establish highly secure communication channels, with a reasonable degree of non-repudiation properties and trust between the different actors involved in the communication.

3.1.5 XML Schema

In order to create the right semantics for interoperability, between different nodes with different implementations, it is essential to have an efficient and highly expressive semantic model for digital identity. This process is highly complex and still requires more comprehensive research from the community as a whole to reach a state where it becomes more practical to automatically reason with identity attributes. [62]

Despite the importance in establishing efficient semantic models for digital identity, this chapter focus is on identity attribute aggregation, so the more complex process involved with digital identity semantics will be addressed as future work. Meanwhile we have designed a more simplified digital identity data representation, based on a XML Schema, that we employ throughout our implementation to keep and promote interoperability for data exchange within OFELIA. The figure 3.2 shows the designed XML Schema skeletal structure that consists in a root element named *OfeliaDataExchange* and it is composed by three main elements: Header, User and Data.

The Header element has two attributes: the State used to describe the current operation and the Type to define the actual stage of the operation. The State operations

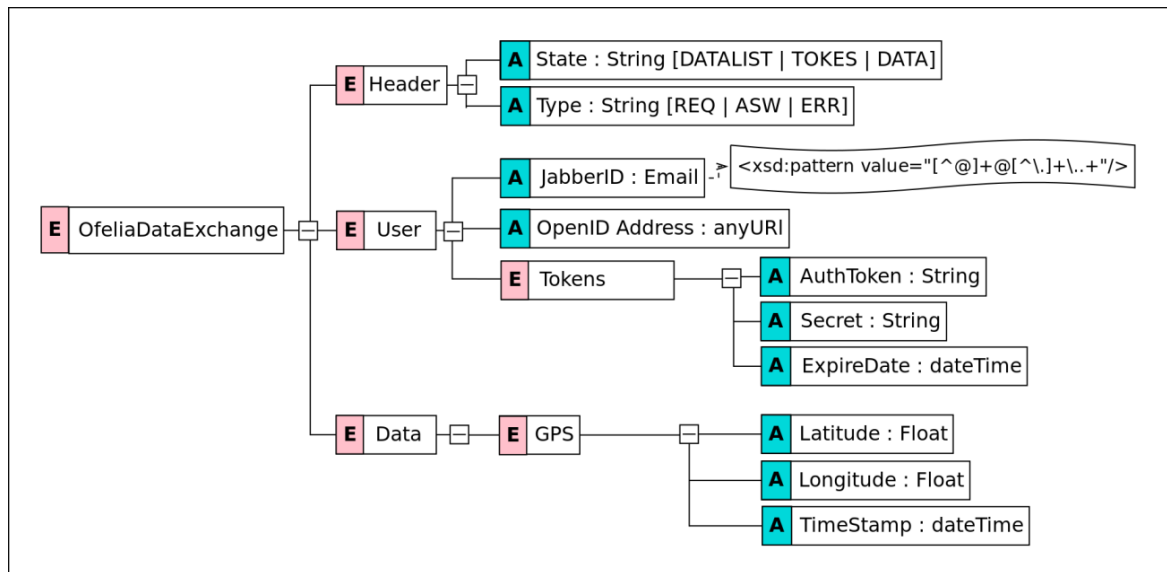


Figure 3.2: Architecture data exchange XML Schema

are classified as: (1) DATALIST used to exchange the list of existing attributes between the smartphone and the identity broker; (2) TOKENS to handle the process of authorization token request; and (3) DATA used to process the data request when data access was previously conceived. The Type is defined in 3 stages: REQ, ASW and ERR that represent respectively request, answer and error.

The User element is composed by three attributes and one element. The attributes are: the JabberID to hold the requester XMPP contact; the OpenID to hold the requester OpenID address and the PubKey to hold requester public key. The element is named Tokens and is composed by three attributes: the AuthToken that is responsible to hold the authorization token; the Secret that acts as a nonce [63]; and the ExpireDate as its own name suggests holds the token expire date.

The Data element is composed by optional elements. Currently we have a gps element defined with the following attributes: Latitude, Longitude and timestamp. We are currently defining several other elements to describe other dynamic attributes like heart beat, blood pressure, etc that could prove to be useful for remote monitoring web applications. The Data element can thus contain highly diverse types of formalised dynamic data types, to cover a Highly diverse range of application areas. In other words, we can provide for all kind of personal dynamic attributes so long as its data type is formalised in the OfeliaDataExchange XML Schema. It is also mandatory that all Data elements have a valid timestamp attribute, not only to be able to maintain an historic value for its values but also to prevent the resending of the same value

during different data exchanges.

3.1.6 QR Code

In our architecture, QR codes are displayed at computers displays to expedite in a secure way the enrollment process of smartphones into the Identity Broker and the Attribute Authorities. QR codes are a very convenient way of conveying a reasonably amount of secret shared information to a smartphone that would otherwise be extremely cumbersome to input by hand by the user.

3.2 Architecture nodes overview

In this section we described the architecture nodes and their importance in our proposed architecture.

3.2.1 OpenID Connect Identity Services

In OFELIA we employ OpenID as an authenticator and as the provider of the bootstrapping information required by the Relying Party to enroll into the Identity Broker. The users essential information that is needed for bootstrapping consists in two key identity attributes, the Identity Broker Internet domain address and the user's public key.

3.2.2 Relying Party/Service Provider (RP/SP)

The Relying party or Service Provider are a web applications that requires users' identity attributes that are being held by the user AAs aggregation. We plan to develop and implement RP/SP software library components to allow for a much more simple integration of current existing web application into our proposed infrastructure.

The software library components must provide functionalities for X509/PGP certificate management, support OpenID Connect authentication and be capable of asynchronously, discover, request, access and store users identity attributes and securely manage authorization tokens. These are issued by the users smartphone, at the users discretion, whenever a RP/SP asks authorization to access a set of users

identity attributes. They contain, among other elements, validity semantic assertions determined by the user that must hold true when the requesting RP/SP presents it to an AA as proof of access entitlement. These tokens are digitally signed by the user at the smartphone to guarantee their integrity and authenticity. An RP/SP must also be capable of secure crypto session keys negotiation with the users AAs by using the IdB as a relay. It must also provide encryption/decryption functionalities for sensitive identity attributes and be capable of parsing and analysing AAs identity assertions according to a digital identity XML semantic specifications. The RP/SP should also provide safe caching of authorization tokens while their validity assertions holds true.

3.2.3 Attribute Authorities

The Attribute Authorities (AAs) are independent network entities responsible for the security and management of personal data. The user smartphone needs to be enrolled into each one of the AAs in order to establish the data aggregation. In order to determine which personal attributes are being held at the AA, the user smartphone is provided with a XML semantic description of the identity attributes that are being held at each enrolled AA. The smartphone then merges the description of the AA identity resources into the users personal data aggregation and announces to the identity broker that it is the custodian aggregator for that data and is now ready to act as a personal authorization broker and issue authorization tokens at the users discretion.

The participating AAs must also be provided with appropriate security mechanisms for authentication and authorization to ensure the appropriate level of access control necessary to protect these assets from unauthorized access and provide the RP/SP with the means to search for identity attributes and negotiate with the identity brokers and user smartphone the authorization tokens needed to be able to access the resources being held by the users aggregation.

This type of framework allows for a simple and scalable integration of an already existing infrastructure of personal data repositories as AAs. For authentication reasons each participating AA must be provided with a public key pair whose authenticity must be attested by a valid PKI X509 or PGP certificate containing the AAs identity. Each AA must also store a list of the emitted authorization tokens whose validity assertions still hold true but have been for some reason revoked by the user.

3.2.4 The Identity Broker

The identity broker acts like a privacy enhancing blind caching-proxy for identity attributes that hides from the RP/SP the real network location of the attribute authority responsible for that data. We need to keep in mind the importance of catering for the situations where the RP/SP cannot be fully trusted and it is therefore important to hide the attribute authority real network location behind a trusted identity broker. Moreover for privacy and security reasons the identity broker must also not know the content of the personal data it is relaying. This is accomplished by having the RP/SP and attribute authority to negotiate session keys and then encrypt all personal data that is being relayed by the identity broker.

The proposed architecture aims for a trust balance where the RP/SP does not have to know about the aggregation of attribute authorities and the identity broker does not need to know about the nature and value of the personal attributes being requested by the RP/SP. For authentication purposes and to prevent men in the middle attacks it is mandatory for the identity broker to be in the possession of a public key pair whose legitimacy can be attested by a valid PKI X509/PGP certificate with the identity brokers' identity.

3.2.5 The smartphone as a Secure Digital Wallet

In OFELIA we are employing android smartphones as highly decentralized personal access authorization management devices for identity management, empowering the user by allowing the creation of customized access control policies that the user finds most adequate for his own personal data. This means that the user is no longer obliged to comply with the abusive identity management policies, normally in place at major sites where the user have to share or give full control of his data to network entities he does not fully know or does not fully trust, as happens with the majority of current Internet applications. OFELIA also brings some advantages in security due to the full hidden decentralization it imposes on the storage of identity attributes.

This application is the critical component of the user digital identity access and should thus be always reachable over the Internet. Unfortunately this is not always possible. Network aware smartphone application are highly demanding in terms of phone battery and network signal usage and therefore cannot be always left running. In order to circumvent this problem the identity broker can be configured by the user

to send a SMS message requesting the smartphone to reconnect. This is archived by the SMS handler service installed on the smartphone in the same time the application is installed. When the SMS handler receives a reconnect SMS message, it launches our application thus reconnecting the smartphone. After a certain period of inactivity our application terminate itself to save on phone battery.

All mechanisms related to authorization token creation, token revocation, attribute access authorization and the enrollment into attribute authorities and the identity broker are conducted by the user interacting with smartphone application. More details about tokens authorization and attribute authorities and identity broker enrollment process are discussed in chapter 4.

Chapter 4

Entities enrollment and case scenario

In our architecture the smartphone plays a key role by acting as the user personal authorization broker. The user starts by enrolling his smartphone into each one of the aggregation participating Attribute Authorities that manage the users personal data. This process allows the mobile device to create an aggregated list of all possible identity data attributes available for that particular user. This list remains solely within the local province of each user personal mobile device and is not disclosed to the network. This helps prevents the massive aggregation of personal data by the Internet operators and gives back to the user some degree of control over his identity attributes.

The smartphone must also be enrolled into an Identity Broker so that the user can then announce and manage the list of attributes names and respective types that can then be made available to the requesting Relying Parties (RP). The authorization tokens needed to access the attributes that are being maintained within the AAs are issued by the users smartphone at the users discretion, after an access request is made by some RP. The creation of the available attributes list is dynamic and must thus be updated each time the smartphone is enrolled or unrolled from an AA, thus increasing or decreasing the number of attributes announced by the identity broker for relying parties, all this under the strict control of the user. In this section we provide a detailed explanation of the different kind of enrollments in a step by step fashion, in order to allow for a better and more comprehensive understanding of the main features provided by the proposed architecture.

4.1 Attribute Authority enrollment

In order to start managing access to his identity attributes, the user first needs to enroll his smartphone with each one of the participating AAs. This process can be done at any time, and should be as effortless and automatic as possible, giving more freedom to the user to painless add or remove AAs as he so wishes. All participating AAs must therefore be OFELIA ready, in other words they must use the AA OFELIA framework and API (mentioned in section 4.1) to properly engage with the other infrastructure participants.

Our architecture provides AAs with an easy and secure method to help the user link his smartphone to the AAs accounts that make up the users attribute aggregation. This is achieved with the help of a specially built AA enrolling web page, where the set of parameters that must be provided to the smartphone to instantiate the linkage with the AA is codified into a specially built QR-code that is displayed on the computer screen as part of the users AA web session. This QR-code is then conveyed to the smartphone by the means of its digital camera. It provides all the necessary URL locations, the AA X509 certificate and the access token the smartphone needs to instantiate the linkage with the AA in a secure way. To enroll is phone with a particular AA the user only has to start an authenticated web session with this AA and then use his smartphone to scan the web session QR-code that is displayed for the enrollment process with the OFELIA application that has already been previously installed in the users personal device. Notice that the smartphone acts as the root node of the AAs, aggregating the user's AAs including their different credentials when applicable.

Figure 4.1 exemplifies the AA enrolment process providing a more technically detailed description of the whole process. This enrollment process is completed in 6 steps:

1. User requests authorization by sending the necessary credential using a web browser.
2. The Attribute Authority grants access if the user credentials are valid.
3. The user request a full access token in order to establish a data access link for the smartphone.
4. The Attribute Authority answers with a data access token and the AA access web services addresses encrypted with the users' public key, all compiled and encoded as a QR code.

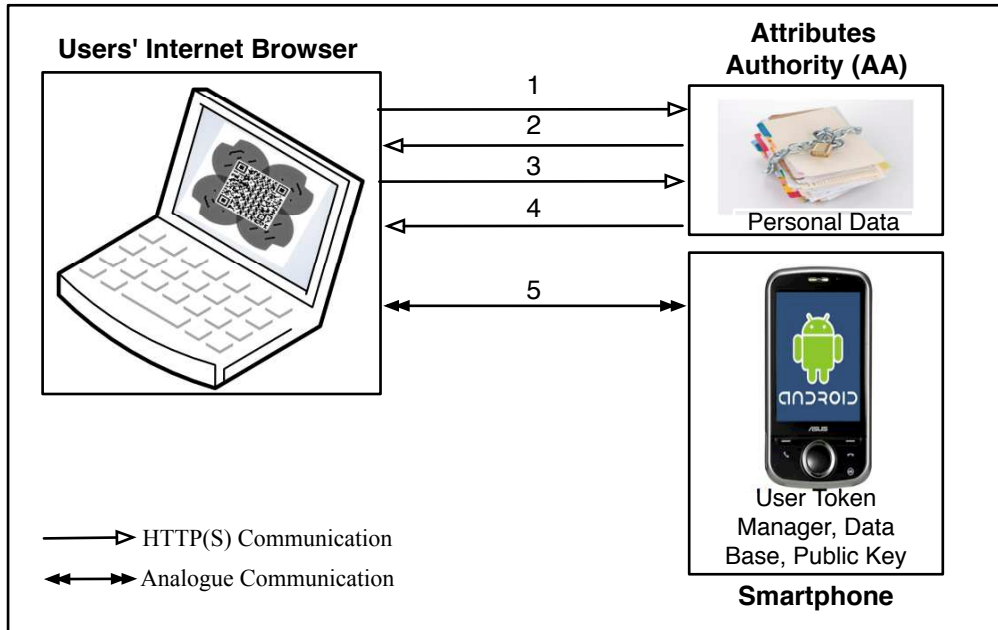


Figure 4.1: AA enrollment flow

- The user uses the OFELIA application in his smartphone to scan the QR-code from the computer screen. The OFELIA mobile application will then automatically proceed and finalize the enrolment process without the need of any further help from the user.

4.2 Identity broker enrollment

In order to establish a communication channel between the relying parties and the user attributes stored in the AAs, the user must also have his smartphone enrolled with an OFELIA identity broker.

This enrollment process between the users smartphone and the identity broker is very similar to the enrollment process described for the AAs. But first the user must use an internet browser to login/authenticate into the IdB with an OpenID Connect account, that is responsible to provides the IdB with an XMPP identity (jabber address) and the public key of the users smartphone. The user is then presented with a QR-code at the computer screen, that can then be scanned by its smartphone, using the OFELIA App. This QR-code contains all the information the smartphone needs to automatically enroll into the IdB. The IdB also provides the user with a web interface

where he can list the history of all the RP/SP attribute requests interactions that have been performed by other third parties. This enrollment process is demonstrated on figure 4.2.

After completing the IdB enrollment process, the user is then free to interact with the mobile OFELIA application to decide upon and determine the restrictions that should be associated with each access requests being made by third party RP/SP web applications. The user can always use the OFELIA App to revoke previously issued and still valid authorizations tokens.

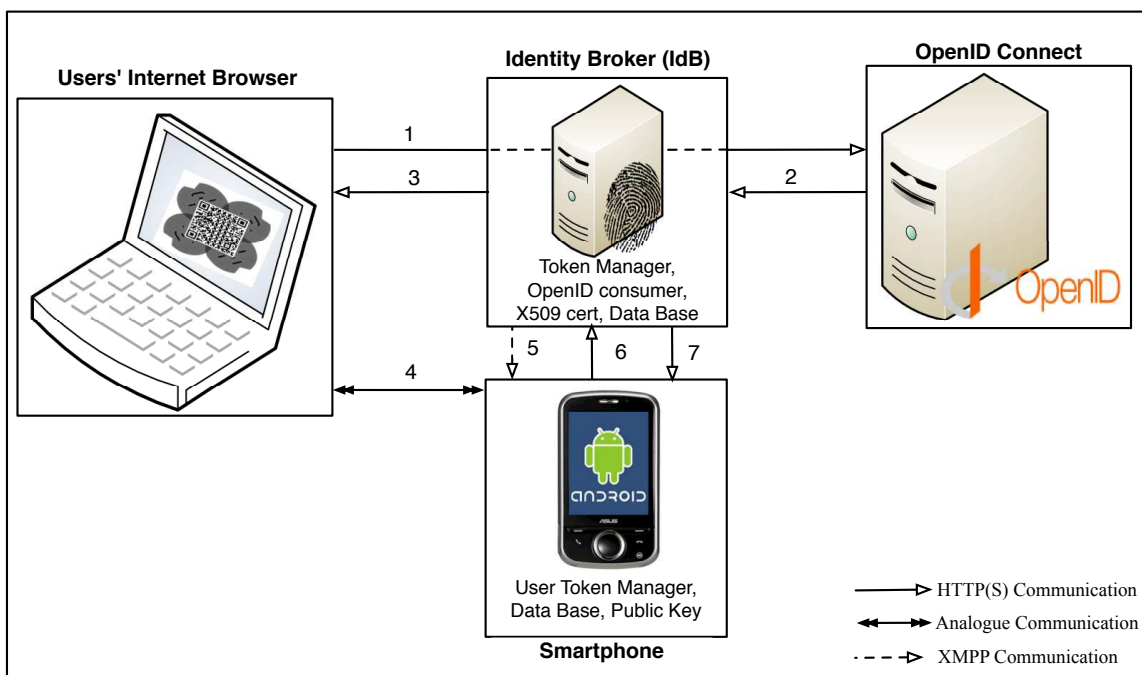


Figure 4.2: Identity Broker enrollment flow

1. The user authenticates at IdB via Openid Connect account and allows the IdB to request the user XMPP address and its public key.
2. The Openid Connect answers to the IdB with the requested data.
3. The IdB sends back to the user computer screen an image of a QR-code of a temporary random link to the IdB session enrollment required data: X509 Certificate, users identification and IdB addresses (XMPP and web addresses) For security reasons this link can only be used once and his discarded by the IdB immediately after use.

4. The OFELIA App scans the QR-code, obtains the link and uses it to retrieve the enrollment data directly from the Internet.
5. The IdB sends to the smartphone an XMPP signed challenge, encrypted with the smartphone public key that has been previously obtained by OpenID Connect.
6. The OFELIA App on the smartphone answers the IdB challenge by sending an XMPP reply containing the list of the attribute names and respective types, that the user desires to share. Note that list is defined by the user in the OFELIA App, if the user do not select the attributes the OFELIA App replies with all the attribute names and respective types that are being aggregated by the users smartphone.
7. The IdB confirms the registration to the users smartphone and this concludes the mobile phone IdB enrollment process.

4.3 Service provider enrollment

Every time the user decides to register a new RP, another enrollment process is triggered in order to allow for the OFELIA requests and data exchange to take place. This process is a bit longer than the other enrollments since we have the participation all OFELIA components.

The user employs an internet browser to logins/authenticates into the RP with its OpenID Connect account, which provides the IdB address as part of one of the users' identity attributes and allows the RP to enroll with IdB as a users' authorized RP application that can ask for the values of a subset of identity attributes approved for that particular RP. After enrollment the RP can then request to the IdB a list of personal attributes. This is done via a XMPP message from the RP to the IdB requesting the list of the available users' data for that RP. This triggers an authorization request from the IdB to the users' smartphone that must be acted upon by the user and leads to the issuing of authorization tokens by the smart phone.

On the users' approval, the OFELIA mobile application creates signed access tokens for each one of the involved data storages (AAs) and also sends an encrypted copy of these access tokens to the RP via the IdB. In this case the encryption is done with the RP public key. This prevents a malicious IdB from issuing data requests on its own. Now the RP can request attributes from IdB while the authorization given by

the user remains valid. This scenario is exemplified in figure 4.3.

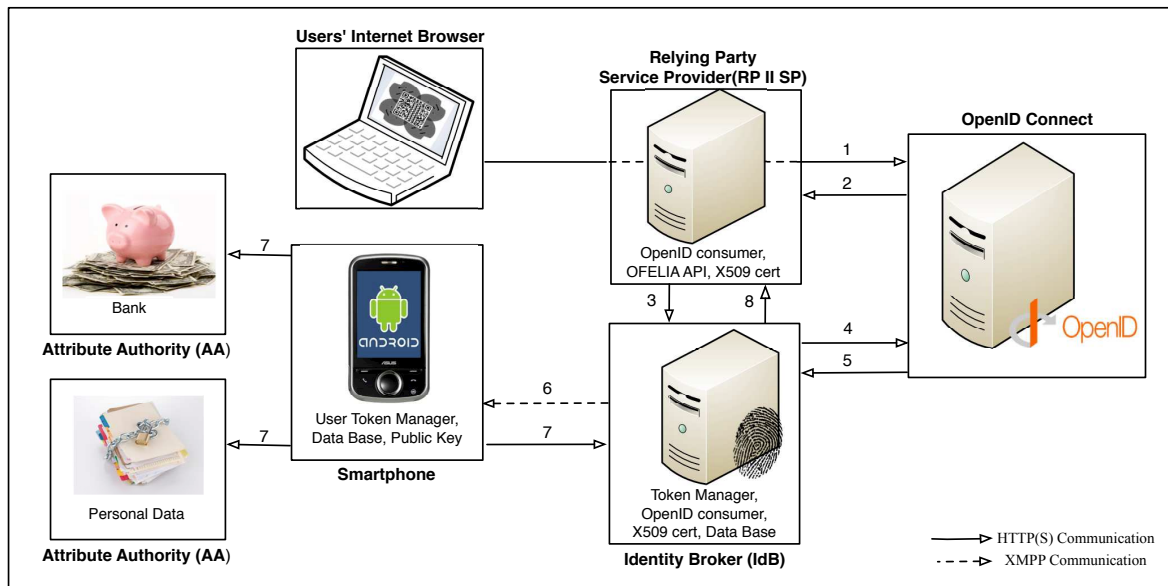


Figure 4.3: RP/SP enrollment flow

1. The user authenticates to the RP via its Openid Connect account allowing the RP to request his public key and the IdB URL HTTP location.
2. Openid Connect answers to the RP with the requested data.
3. The RP makes a TLS REST registration request to the IdB, providing its certificate as the client cert for the TLS connection that is established from the RP to the IdB. The registration request contains the OpenID request link, some descriptive information details (to be displayed at the users mobile phone) about the RP service and a list of the requested attributes encrypted by the user's public key.
4. The IdB tests the OpenID request link in order to verify if the request is valid.
5. Openid Connect answers the IdB, If the answer from the OpenID server comes as a replay-attack [63] attempt, it in fact confirms to the IdB that the user has been previously authenticate with OpenID at the requesting RP and therefore this RP enrollment attempt is legitimate. This is a widespread OpenID hack that allows a service to verify if the user has already been previously OpenId authenticated at some other site. The IdB can then pre-register the RP by generating an RP identifier token.

6. The IdB sends a XMPP message to the smartphone containing a signed request message with the encrypted RP data request plus other requesting RP details (identifier token, certificate, details of service and RP URL HTTP location).
7. At the users' discretion, an access authorization token is generated by the smartphone and sent back to the IdB encrypted with the RP public key and encrypted with each AA public key to each one of the involved AA with the RP details.
8. The IdB validates the RP registration by sending to the RP the encrypted access token that has been issued by the smartphone.

4.4 Usage case Scenario

In this section we presented two distinct case scenarios of our architecture. The first focused on the attribute aggregation paradigm and its benefits on e-commerce [64]. The second case scenario focused on the attribute authorization access problem and its advantages on e-health sector.

4.4.1 Attribute Aggregation scenario in e-commerce

For a credible illustrative OFELIA aggregation scenario, imagine an online bookstore as a Relying Party and for example a credit card company and university acting as Attribute Authorities. Now lets assume the user is online shopping at the online bookstore and upon completion of his purchase, if he can prove that he has a specific bank card and is a student of certain university, the online bookstore gives him an immediate special discount on books of his study domain.

At the moment of purchase and after the user had already been authenticated via OpenID Connect, the online bookstore, acting as a RP, will request the IdB of that user for proof of bank card and university membership for that particular user. This triggers an authorization request made by the IdB that is displayed at the user smartphone, to authorize the necessary AAs to disclose this information. The user can then use the OFELIA App application installed at his smartphone to authorize both AAs (university and bank card) to disclose the users membership status (signed by the AAs X509 certificates) to the bookstore. These authorizations take the form of digitally signed authorization tokens that are registered on the respective AAs and

delivered to the IdB encrypted with the RP public keys. The IdB then acts as a relay and sends the signed encrypted authorization tokens back to the Relying online bookstore (RP).

The RP, now in possession of these digitally signed authorization tokens, can then send them to the IdB, encrypted with the respective AA public key each time the online bookstore wants to get evidence the user is still a valid customer of the bank and member of an university. These access tokens together with the identity consultation requests are then digitally signed and relayed by the IdB into the appropriate AAs, which upon analyzing the validity of the accompanying authorization tokens can deliver the requested information back to the IdB, digitally signed by the AAs and encrypted for the RP. This encryption step is important in order to establish a high level of privacy and security. The IdB should not know the value of the identity attributes, otherwise the entity responsible for the IdB would be in a position of doing massive data aggregation with their users data, and that aggregation by itself would become a much more prized target for attacks. This constitutes two of the main reasons for OFELIA to have been developed in the first place, i.e, to provide an identity/authorization versatile infrastructure that does not depend upon the massive aggregation of users' identity attributes.

Finally the IdB relays the requested encrypted information to the RP that can verify its integrity and validity by decrypting the attributes values and verifying the validity of its digital signatures and thus letting the online bookstore (RP) apply the special discount on books of the buyer subject studies domains.

4.4.2 Attribute authorization access scenario in healthcare

To better understand our architecture capabilities on e-health sector we defined a storyboard of a patients' parent, more precisely, a daughter that wants to access her mothers' (the patient) medical information by the means of her computer.

Katherine, a 50 years old woman that resides in Mystic Falls, has recently finished her radiotherapy treatments after being diagnosed with breast cancer. Her daughter, Agnes, who lives in a different city 600km away, desires to monitor her mothers follow-up consultations. Assuming that both mother and daughter are already enrolled at the same healthcare institution, Agnes, the daughter, requires Katherine, the EHR owner, an authorization to access Katherines' EHR with her smartphone. Katherine, also using her smartphone decides to grant access to her daughter. However, Katherine wants

to customize some of the access control rules since she desires to omit the treatments record component, creating the specific role Patients Daughter for that purpose. Now Agnes accesses her healthcare institution website (that triggers the adapted version of our identity broker) with her browser and a QR code is returned and read by the smartphone application that handles the authorization process with the XMPP server into the adapted identity broker. After that, Agnes browser automatically refreshes with a list of the patients for which she has permissions to access. Now Agnes selects her mother assuming her assigned role, Patients Daughter, allowing Agnes to read the wished follow-up consultations record component.

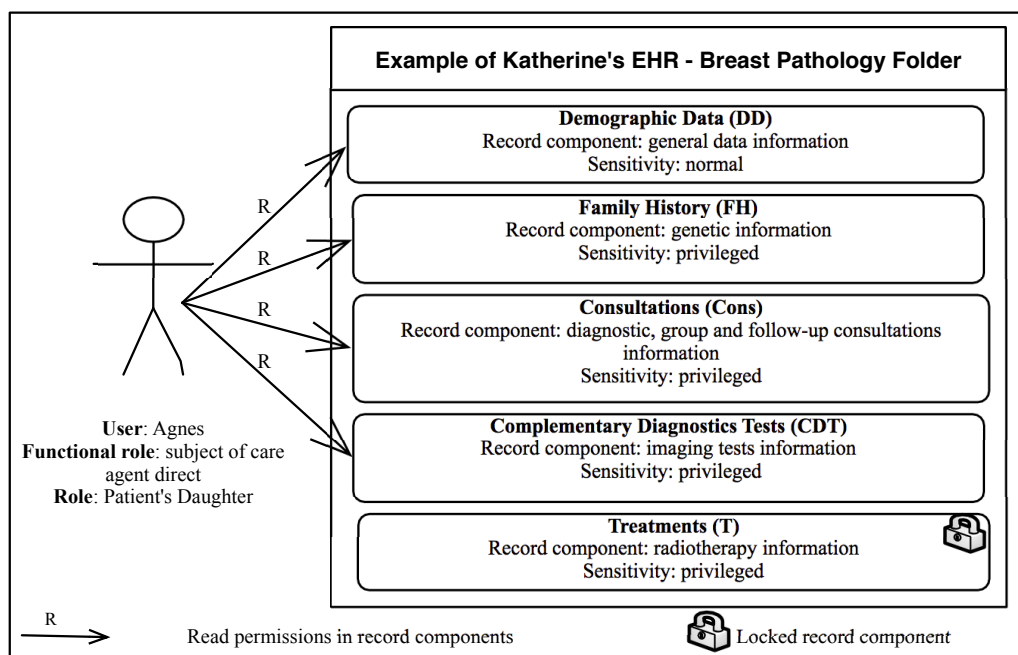


Figure 4.4: Agnes access permissions to Katherine's breast pathology folder

Figure 4.4 illustrates the use case of the above described storyboard. This use-case shows an example of the EHR folder regarding the Breast Pathology components [65]. The user Agnes accesses Katherine's EHR, with the Patients Daughter role, attributed by her mother, which gives Agnes permissions to only read the following components: demographic data, family history, consultations and complementary diagnostics tests. Due to the role restrictions made by her mother, Agnes cannot access the treatments component.

Chapter 5

Conclusions and Future Work

We are currently sitting at a crossroads where the mega Internet corporations of today are amassing an inordinate amount of personal information and thus concentrating too much power into their own hands. These Internet giants are comfortably embarking on a journey that is certainly brilliant for their shareholders but very bleak and unbalanced for the common user, who is most of the time unaware that he is the product being sold at the services he so diligently uses on the Internet. The service providers feel very comfortable with this unbalanced situation and they will not make any effort to change this status quo. But the future is yet to be written and if we, the users, are given back some level control about our personal data, we will be able to transport ourselves into more joyous shores and reach other destinations where users are treated not simply as merchandise to be sold on the Internet but as real business partners that only disclose their most sensitive identity attributes when needed and when the “price” is right. However this market for personal data attributes can only be fair, and provide the user with a fair share of the profits, if he is provided with the technical means to disclose personal data for limited periods of time and keep those authorizations under strict revocation control. We firmly believe OFELIA constitutes a firm step in the right direction, and hope to have convinced the reader, with the work we have presented with this document, that this is so.

5.1 Research summary

We have studied the main identity management and attribute aggregation models to identify and take inspiration from the best of breed practices. We have then integrated

these into an innovative user-centric mobile device based system. This system can be used by users to self-manage access control to aggregate identity attributes in a versatile, distributed, secure and privacy enhanced way. At the same time users are provided with the means to customize on the fly revocable access control assertions that can be fully adapted to the entity that is requesting access to personal data. The OFELIA framework architecture is the result embodiment of these requirements.

With this dissertation we have described a fully working user centric identity management and attribute aggregation linking model, enhanced with mobile devices and provided with innovative solutions for facilitating all the entities enrollment into the identity management infrastructure. We have also described and taken inspiration from several other proposed identity management systems, because they serve to illustrate and help explain very useful security mechanisms and communications protocols (such as OpenID and OAuth) that have been directly or indirectly integrated into the OFELIA architecture. These are: (1) XMPP, RESTful web services and QR-Codes for communications; (2) Public Key Infrastructure (including PGP and X509 certificates), smartcards and authorization tokens to ensure security and privacy; (3) XML Schemas to guarantee interoperability between the different architectural components.

We believe we have accomplished the main objectives that we have propose to achieve at the beginning of the work that lead to this dissertation. However OFELIA is very much a work in progress and what has been done until now can only be seen as a first step for a more complete a proof of concept that must now be exercised by being evaluated in the context of real usage case scenarios.

5.2 Main findings

We can infer from the current state of the art (Chapter 2) that identity management and attribute aggregation models are nowadays very hot research topics with a high interest not only due to the sheer number of authors that dedicate their research budgets over it, but also by the the real impact they can have on society and the way hundreds of millions of users will interact with the services provided by the Internet in the near future. However we have also found that none of the IdMs we have studied so far fully satisfied on their own our initial research goals and questions about secure user-centric and distributed identity attribute aggregation. Nevertheless they provided us with security mechanisms and guidelines that where fundamental to integrate our ideas into a new user-centric mobile based architecture embodying the research goals

we had set for ourselves at the beginning of our work.

Our identity attribute aggregation architecture allows for a greater participation, responsibility and control over who can access personal information by the user. In OFELIA the user is responsible for the distribution of his identity attributes over the Internet and have them managed and controlled by fully integrated Attribute Authorities. These attribute authorities should be the direct source and final authority about the identity attribute. This allow us to also refine the concept of highly dynamic identity attributes, by having AAs directly attached to the data-source. The user can also only aggregate the necessary identity attributes when needed and this allow us to embody the highly popular concept of the identity persona.

Access to the users' identity attributes is restricted only to pre-authorized requesters (service providers or users). These authorizations are emitted by the users' smartphone at the user discretion and take the form of authorization tokens with temporal constrains and access rules. We have argued and hope to have convinced the reader about the merits of using the smartphone as the "de facto" mobile asynchronous authorization broker.

XMPP has also proved to be an excellent choice for the establishment of secure and asynchronous communication channels between the smartphone, the identity broker and the attribute authorities. Its benefits were already described in section 2.5.1 and have been experimented first hand during the implementation work we have conducted for the OFELIA proof of concept.

5.3 Current OFELIA implementation limitations

In order to take advantage of OFELIA the user has to understand and possess basic skills on the use information technologies(IT) and be basically proficient on the use of smartphones. Other limitation is that the mobile application has currently only been developed for the Android platform, which restricts the scope of its users.

The use of a non-standard identity XML specification language can be seen as problem for the architecture interoperability and acceptability. The XML Schema we have used have only been created for interoperability between the architecture nodes, since identity modelling was not the focus of this thesis and the use of other XML languages like SAML or XACML, could affect our choices over the architecture due to their already made profiles and their specific library implementations. However we recognize

the high importance of appropriate identity modelling and intend to address it as future work.

Due to the highly divergence of data providers and services providers, the full integration with already running identity providers (Google, Facebook) is not as well supported in OFELIA as we would like it to be. This will be addressed as future work as well.

5.4 Future work

Portability is essential in order to maximize the interest in our architecture. For now the smartphone application has been developed for the android platform. We want to port it to other smartphone systems like IOS, BADA and Windows phone.

We also want to concentrate our future research efforts into the construction of a stronger semantic models for identity management. This is highly complex but necessary for services where automated management and reasoning with personal attributes could be useful. At the moment OFELIA has a very simple semantics for identity attributes that work as a starting proof of concept. A research about identity semantic models and the adoptions of standards such as SAML are currently being addressed as future plans.

Digital signature and encryption are also fundamental parts of our proposed architecture, but at the moment they are yet to be integrated and are in development due to the fact that we are still running tests with mobile security cards [55] and figuring the best way to integrate them into the Android mobile platform.

We are also currently implementing and deploying OFELIA in the health sector [22] with a very specific user-empowering usage case scenario in a real healthcare institution, more precisely on São João hospital centre, which is the second biggest hospital in Portugal.

5.5 Conclusion

The proposed architecture is an user centric empowering infrastructure where it is possible to securely dynamically manage the aggregation of identity attributes from different authorization authorities into a single user centric digital identity whose

authorizations can be managed in a novel versatile way involving for example temporal constraints by the arbitrage of the user smartphone. An versatile infrastructure like this can be easily applied in different application domains.

The architecture also possesses innovative mechanisms to protect users privacy by preventing the massive aggregation of users identity attributes into a single place. We have taken special care to prevent the disclosure of identity attributes values at the identity broker precisely to prevent the massive disclosure of user data lest the identity broker be compromised. Thus if an attacker compromises the identity broker he will not have disclosed the users identity attributes values that should therefore continue to remain safe in a privacy aware away. Furthermore since the identity attributes are always held by their original source (the attribute authority) the identity attributes maintains a kind of freshness state.

Appendix A

Development notes

We have employed several open source libraries and applications to implement the different OFELIA components and assemble a testbed to run demo OFELIA applications. We have developed a demo web application (relying party) where a user authorizes another to track his GPS positioning for a certain period of time. The Relying party application uses the OFELIA infrastructure to obtain authorization tokens from the users mobile phone and then uses these tokens to obtain the users GPS coordinates from the users mobile device (that is running an AA implementation that knows about the phone GPS). The application then maps the GPS coordinates for visualization by using the services provided by google maps for the requesting user to see.

In this appendix we review the software libraries we have used to assemble the OFELIA testbed and develop the fully functional demo we have just described.

We also presented some real application screen captures.

A.1 Identity Broker Web Service

The identity broker is a RESTful web service developed in Oracle Java SE 7 by the usage of the Java development kit (JDK) and deployed on the Apache Tomcat Servlet/JSP container version 7.

The libraries employed for the implementation were:

- **Apache Xerces2 Java** is a library for parsing, validating and manipulating

XML documents.

- Website: <http://xerces.apache.org/xerces2-j/>
- Version: 2.11.0
- Download link: <http://mirrors.fe.up.pt/pub/apache//xerces/j/binaries/Xerces-J-bin.2.11.0.tar.gz>
- **Openid4java** as its own name suggests is a library that provides the right means to consume OpenID identities for Java applications.
 - Website: <http://code.google.com/p/openid4java/>
 - Version: 0.9.5.593
 - Download link: <http://openid4java.googlecode.com/files/openid4java-full-0.9.5.593.tar.gz>
- **Smack API** is a Java open source XMPP client library for instant messaging and presence from Ignite Realtime.
 - Website: <http://www.igniterealtime.org/projects/smack/>
 - Version: 3.2.2
 - Download link: http://www.igniterealtime.org/downloads/download-landing.jsp?file=smack/smack_3_2_2.tar.gz
- **MySQL Connector** offers a standard database driver connectivity for using MySQL with applications and tools that are compatible with industry standards ODBC and JDBC. Since we are developing in Java we used the Connector/J.
 - Website: <http://www.mysql.com/downloads/connector/>
 - Version: 5.1.22
 - Download link: <http://www.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.22.tar.gz/from/http://cdn.mysql.com/>
- **ZXing** is an open-source, multi-format 1D/2D barcode image processing library implemented in Java. In the identity broker is responsible to generate the QR-Codes.
 - Website: <http://code.google.com/p/zxing/>
 - Version: 2.0

– Download link: <http://zxing.googlecode.com/files/ZXing-2.0.zip>

Figure A.1 presents the identity broker package tree that is composed by three packages:

(1) The *org.ofelia* is composed by just one class, the *core.java* responsible to handle and redirect all the identity broker requests (XMPP or HTTP), this class acts as the identity broker backbone.

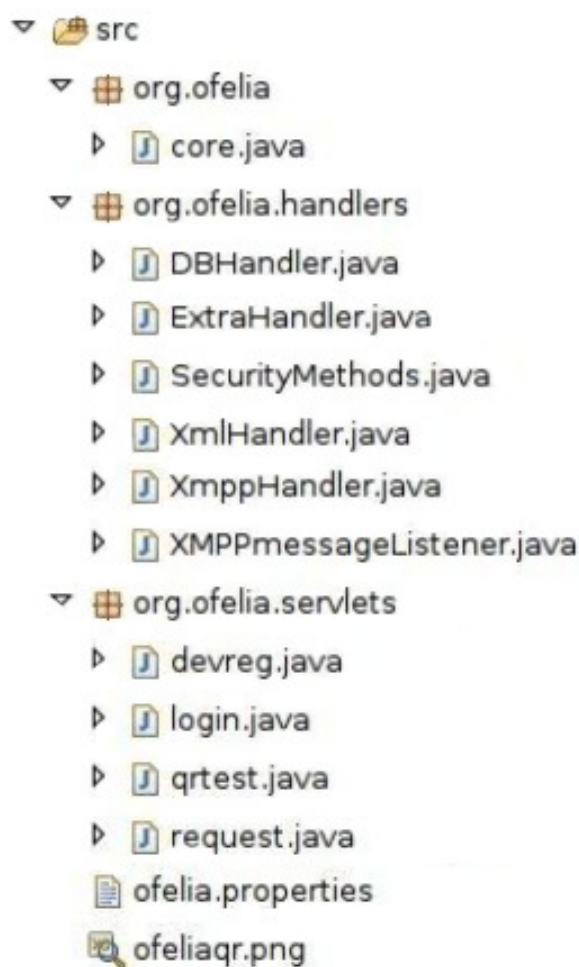


Figure A.1: IdB package tree

devreg.java responsible to register new AAs at the identity broker; (ii) the *login.java* to handle the user registration and login process by the usage of an OpenID server; (iii) the *qrtest.java* is used to generated the necessary QR codes to guarantee the enrollment process; (iv) and the *request.java* responsible to handle the authorization process by preparing and sending an xmpp message to the users' smartphone.

(2) The *org.ofelia.handler* is composed by six classes that act as the web service handlers: (i) The *DBHandler.java* handles the necessary SQL functions to read and write the identity broker database; (ii) the *ExtraHandler.java* is responsible to hold generic functions as simple math or print methods; (iii) the *SecurityMethods.java* handles the functions to manage the access tokens and the cryptographic methods; (iv) the *XmlHandler.java* holds the methods to read, write and validate the XML documents; (v) the *XmppHandler.java* stores the methods to establish and manage a xmpp connections; (vi) and the *XMPPmessageListener.java* interface implements the observer pattern in order to notify the *core* class every time a xmpp message is received.

(3) The *org.ofelia.servlets* is composed by four servlets that are normally invoked by an OFELIA JavaServer Page (JSP). These four servlets are: (i) the

The *ofelia.properties* is the configuration file that holds the XMPP and SQL server accounts and the *ofeliaqr.png* is the logo of the OFELIA project.

Figure A.2 represents the identity broker enhanced entity relationship (EER) database model. This database allows to store the necessary information to sustain our identity broker. In this EER there are two type of possible data to access the *CARDIO* and the *GPS*. Both of them have the *TStamp* (time stamp) field to keep a temporal coherence and the *AccuracyLevel* to determine the level of precision of the information that was fulfilled by the attribute authority. (e.g. There are several methods to obtain a GPS coordinate however some methods are more precisely than others). The *DEVICE* table holds the users' attribute authorities and their supported data in the table *DEVICEDATA*.

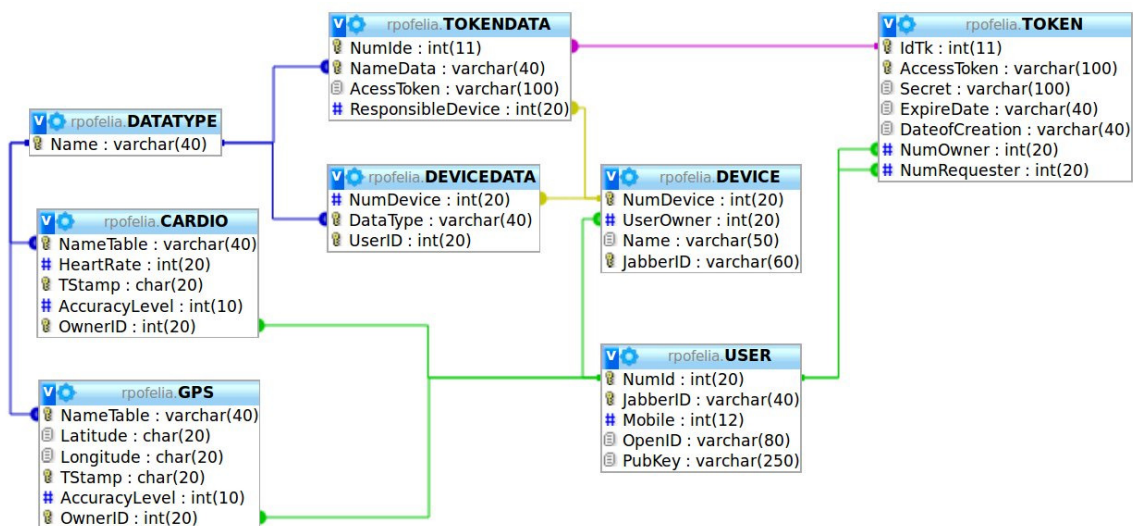


Figure A.2: Identity Broker Database EER

A.2 Attribute Authority standalone application

The Attribute Authority (AA) standalone application was developed on Java SE 6 and ported to Android smartphones in order to transform these devices into AAs.

The libraries employed for the android port were:

- **Android SDK (API 10)** provides the necessary API libraries and developer tools to build, test, and debug applications for Android devices.
 - Website: <http://developer.android.com/sdk/index.html>

- Version: 20.0.3
- Download link: http://dl.google.com/android/android-sdk_r20.0.3-linux.tgz
- **asmack** is a Java open source XMPP client library for instant messaging and presence based on the Smack API from Ignite Realtime for android.
 - Website: <https://code.google.com/p/asmack/>
 - Version: issue15
 - Download link: <https://asmack.googlecode.com/files/asmack-issue15.jar>
- **ZXing** to create the right means for the Attribute Authority android application read the QR codes for their enrollment process.

Figure A.3 shows the AA package tree that is composed by three packages:

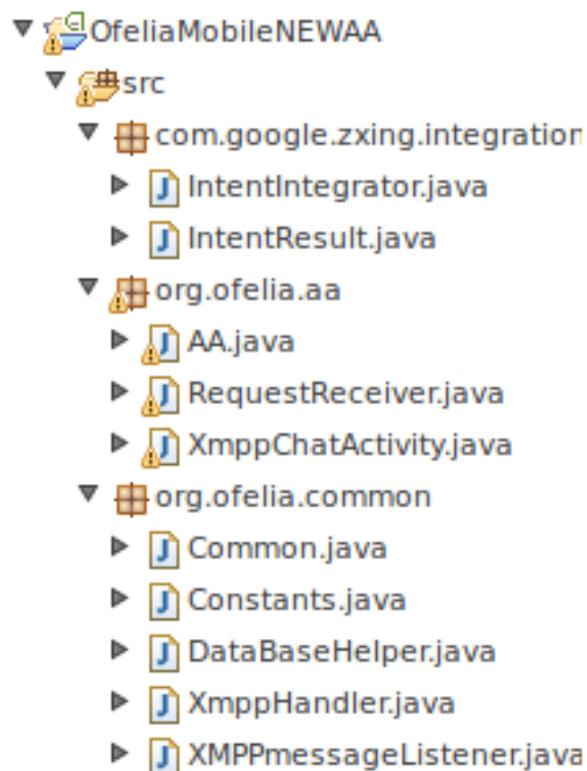


Figure A.3: AA Android package tree

(1) The *com.google.zxing.integrator* is responsible to establish the right means to read the QR codes, this package is part of the ZXing library.

(2) The *org.ofelia.aa* that contains three classes responsible for the AA core implementation: (a) The *AA.java* that holds the main methods for the AA implementation, methods like the token management and data management; (b) the *RequestReceiver.java* implements a service that waits for a specific SMS that requests the smartphone to turn on its internet and the main application when requested; and (c) the *XmppChatActivity.java* that is responsible to construct the mobile program interface and manage the XMPP connections.

(3) The *org.ofelia.common* that contains the common classes used in android every time an application is being developed for OFELIA. This package contains five classes: (a) the *Common.java*, holds the generic methods like math or print methods; (b)

Constants.java like its own name suggests it stores the application program constants; (c) The *DataBaseHelper.java* handles the necessary SQL functions to read and write the AA database; (d,e) the classes *XmppHandler.java* and *XMPPmessageListener.java* work exactly like in the identity broker web service (see A.1).

Figure A.4 represents the Attribute Authority EER database model. This database has several similarities with the identity broker database. The table *USER* is used to represent the requesters.

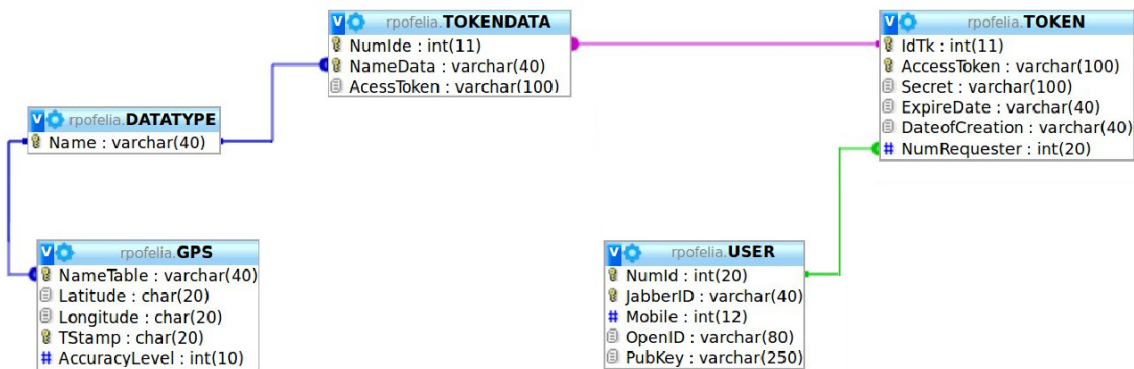


Figure A.4: Attribute Authority Database EER

A.3 Secure Digital Wallet application

The Secure Digital Wallet application is a pure android application.

The libraries employed for this android application were:

- **Android SDK (API 10)** to provides the necessary API libraries and developer tools to developed our application.
- **asmack** used to establish the XMPP connections.
- **ZXing library** in order to consume the QR codes.

Figure A.5 presents the Digital Wallet application package tree. This android application for smartphones is composed by three different packages:

(1,2) The packages *com.google.zxing.inte- gration* and *org.ofelia.common* are the same from the Attribute Authority android application. (see A.3)

(3) The package *org.ofelia.wallet* implements the core functions of this android application and is composed by four classes: (a) The *FirstTimeView.java* is an android activity that appear only once at the first time the user opens the application in order obtain the necessary information to deploy the application (a XMPP account and a PIN);

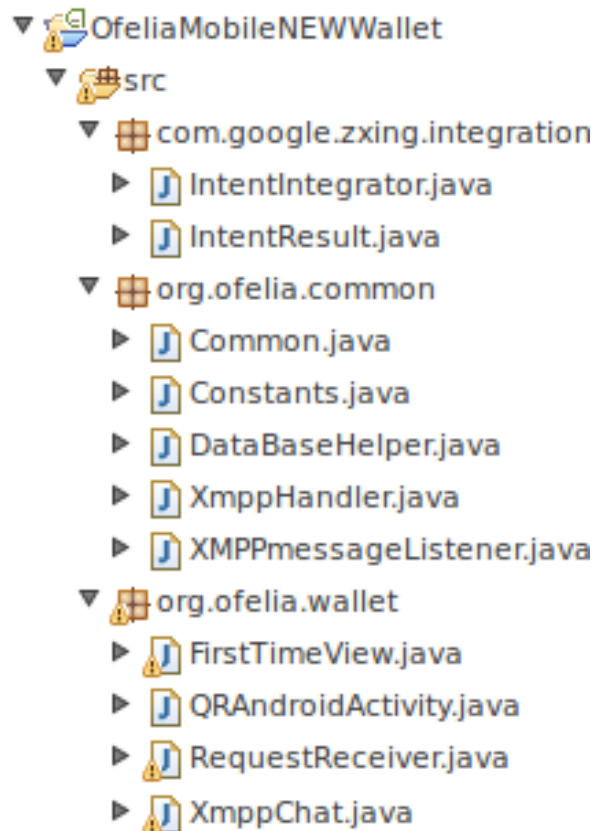


Figure A.5: Digital wallet Android package tree

(b) The *QRAndroidActivity.java* is the android activity that allows the QR code enrollment process; (c) The class *RequestedReceiver.java* implements a SMS service already explained on the Attribute Authority standalone application (see A.3; and (d) the *XmppChat.java* is the main class on this application since defines the interface, establish the communication and handles the process of authorization.

Figure A.6 represents the digital wallet EER database model. This database has several similarities with the identity broker database. However there is no data tables (like GPS) since the digital wallet is not a consumer or producer of data information. The DEVICE table is used to store the users' AAs information.

A.4 Relying Party web service

Despite the fact that we have developed an API for already running relying parties we developed our own Relying Party (RP) web service in order to evaluate our architecture. This web service was developed in Oracle Java SE 7 by the usage of the Java development kit (JDK) and deployed on the Apache Tomcat Servlet/JSP container version 7.

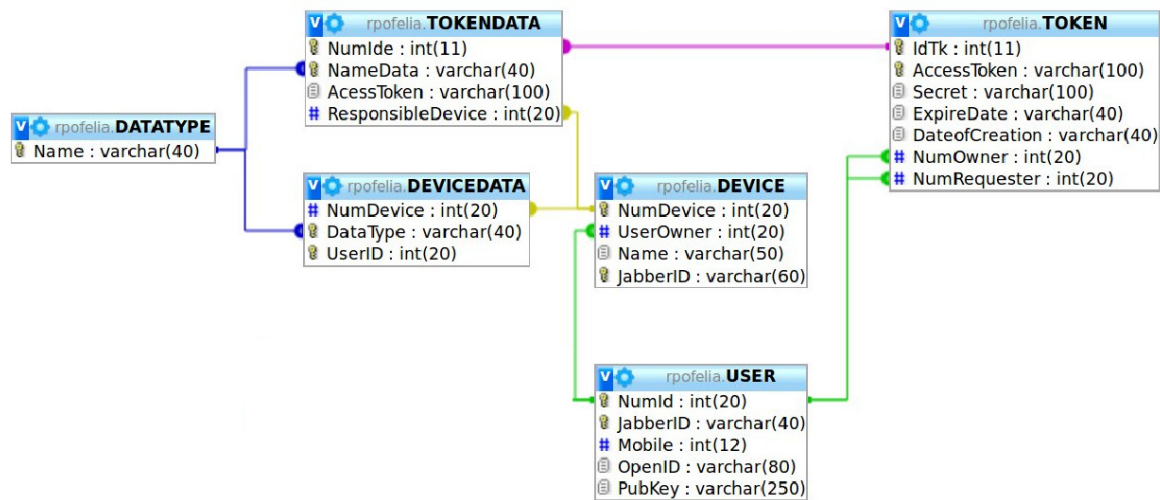


Figure A.6: Secure digital wallet Database EER

The libraries employed for the implementation were:

- **Apache Xerces2 Java** used to parsing, validating and manipulating XML documents.
- **Openid4java** employed as a consumer of OpenID identity
- **MySQL Connector/J** to establish connection between a MySQL database and the web service.



Figure A.7: Relying party package tree

Figure A.7 shows the relying party web service package tree. This figure is formed by three groups of packages:

(1) The *org.ofelia* is composed by just one class, the *core.java* responsible to handle and redirect all the relying parties HTTP requests, this class acts as the RP backbone.

(2) The *org.ofelia.handler* is composed by two classes: (i) The *DBHandler.java* and the *XmlHandler.java* both already described in the Identity Broker (see A.1)

(3) The *org.ofelia.servlets* is composed by three servlets that are invoked by

an OFELIA JavaServer Page (JSP). These servlets are: (i) *dataVisualizer.java* responsible to analysis the data types received in order to give the best data output view for the requester (at the moment only recognizes the GPS data that is represented as coordinates on google maps); (ii) the *login.java* to handle the user registration and login process by the usage of an OpenID server; (iii) and the *requestTokenData* is used to request new data accesses to the Identity Broker.

Figure A.8 represents the Relying Party EER database model. In this database the table *USER* holds information only about the data consumers. The table *TOKEN* and *TOKENDATA* holds information for data consumption such as the type of data, their expire date and the jabber address contact from the requested user.

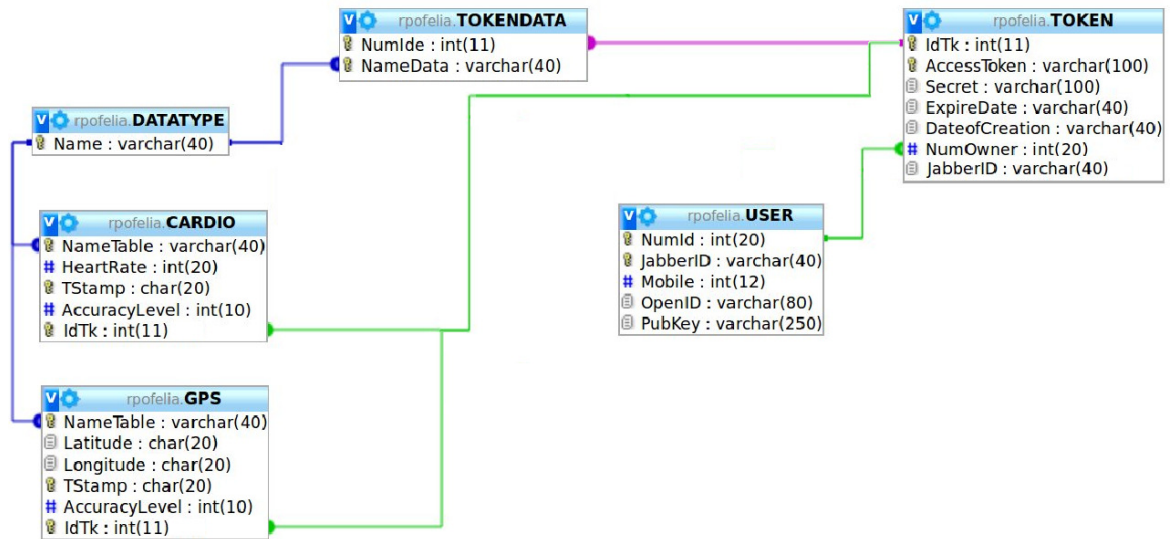


Figure A.8: Relying Party Database EER

A.5 Screen captures:

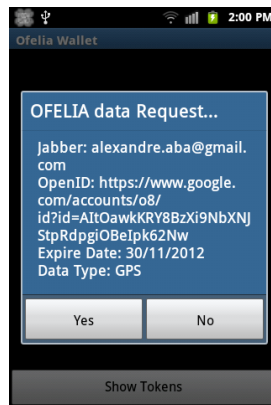


Figure A.9: Ofelia digital wallet authorization request box.

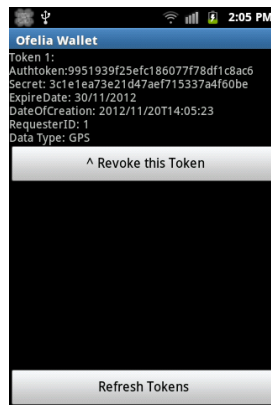


Figure A.10: Ofelia digital wallet token list.

Welcome: alexandre.aba@gmail.com

RP Section:

[Request new Data](#)

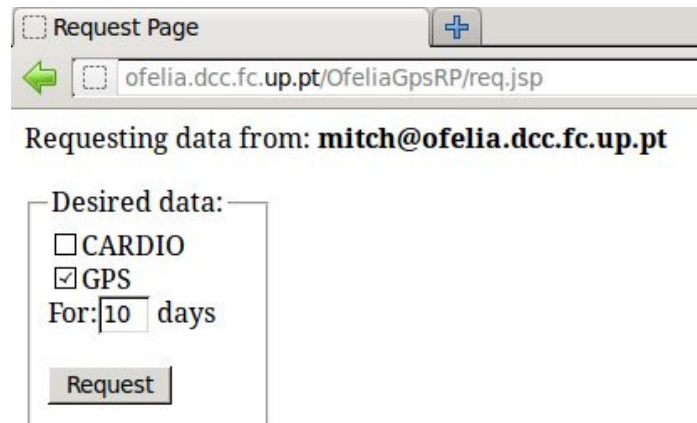
[Pending Tokens \(1\)](#)

[Active Tokens \(1\)](#)

[History](#)

[Logout](#)

Figure A.11: Relying Party user area.



Request Page

ofelia.dcc.fc.up.pt/OfeliaGpsRP/req.jsp

Requesting data from: **mitch@ofelia.dcc.fc.up.pt**

Desired data:

CARDIO

GPS

For: days

Request

Figure A.12: Relying Party data request page.



Figure A.13: Relying Party GPS coordinate consultation.

References

- [1] United Nations. *Universal declaration of human rights, Article 12*. U.S. Govt. Print. Off., Washington, 1949. 1
- [2] David Recordon and Drummond Reed. Openid 2.0: a platform for user-centric identity management. In *Proceedings of the second ACM workshop on Digital identity management*, DIM 06, pages 11–16, New York, NY, USA, 2006. ACM. 2, 20
- [3] E. Hammer-Lahav. The oauth 1.0 protocol (rfc5849). <http://tools.ietf.org/html/rfc5849> Verified on 14/10/2012, April 2010. 2, 34
- [4] Mont Marco; Pearson Siani; Pete Bramhall. Towards accountable management of privacy and identity. *Information: Computer Security, ESORICS; Lecture Notes in Computer Science*, pages 146–161. 2, 7
- [5] K. Tracy. Identity management systems. *Potentials, IEEE*, (2):34–37, Nov 2008. 2
- [6] Hai-Binh Le; Bouzefrane S. Identity management systems and interoperability in a heterogeneous environment. *Advanced Technologies for Communications, 2008. ATC 2008. International Conference*, pages 239–242, Oct 2008. 2, 19
- [7] Alexandre B. Augusto and Manuel Eduardo Correia. An xmpp messaging infrastructure for a mobile held security identity wallet of personal and private dynamic identity attributes. In *Proceedings of the XATA 2011 XML: Aplicações e Tecnologias Associadas*, 2011. 4, 8
- [8] Barkhuus L. and Polichar V. Empowerment through seamfulness: smart phones in everyday life. *Personal and Ubiquitous Computing*, page 629639, 2011. 4, 33

- [9] W. Adi; A. Al-Qayedi; A. Zarooni.; Mabrouk. Secured multi-identity mobile infrastructure and offline mobile-assisted micro-payment application. *Wireless Communications and Networking Conference, 2004. WCNC.*, March 2004. 4
- [10] F. Paci; Ning Shang; K. Steuer; R. Fernando; Bertino. Veryidx - a privacy preserving digital identity management system for mobile devices. *Mobile Data Management: Systems, Services and Middleware, MDM09. Tenth International Conference*, May 2009. 4
- [11] Zhikui Chen. A privacy enabled service authorization based on a user-centric virtual identity management system. *Communications and Networking in China, 2007. CHINACOM 07. Second International Conference*, August 2007. 4
- [12] A. Josang and S. Pope. User-centric identity management. *Proceedings of AusCERT 2005, Brisbane, Australia*, May 2005. 4, 21
- [13] Alexandre B. Augusto and Manuel E. Correia. Ofelia - a secure mobile attribute aggregation infrastructure for user-centric identity management. In *Proceedings of the IFIP SEC2012 (International Information Security and Privacy Conference)*, Creta, Greece, June 2012. Springer IFIP Advances in Information and Communication Technology. 4, 8
- [14] D. Gollmann. Computer security. *Wiley Interdisciplinary Reviews: Computational Statistics, John Wiley & Sons, Inc.*, (2):544554, Jul 2010. 7
- [15] Paul M. Schwartz. Property, Privacy, and Personal Data. *SSRN eLibrary*. 7
- [16] E. Hammer-Lahav. Security architecture. <http://bit.ly/OAuthToken> Verified on 14/10/2012, October 2008. 7
- [17] Eran Hammer-Lahav. Introducing oauth 2.0, 2010. 7
- [18] Nat Sakimura; John Bradley; Breno de Medeiros; Michael Jones; Edmund Jay. Openid connect standard 1.0. <http://tinyurl.com/openidc> Verified on 13/01/2012. 7, 20
- [19] Kristiina Hayrinen, Kaija Saranto, and Pirkko Nykanen. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*, 77(5):291–304, 2008. 8
- [20] Mor Peleg, Dizza Beimel, Dov Dori, and Yaron Denekamp. Situation-based access control: Privacy management via modeling of patient data access scenarios. *Journal of Biomedical Informatics*, 41(6):1028–1040, December 2008. 8

- [21] Dept. of Health & HS. The office of the national coordinator for health information technology, 2011. 8
- [22] Cátia Santos-Pereira, Alexandre B. Augusto, Manuel Eduardo Correia, Ricardo Cruz-Correia, and Ana Ferreira. A mobile based authorization mechanism for patient managed role based access control. In *3rd International Conference on Information Technology in Bio- and Medical Informatics - ITBAM 2012*, Vienna, Austria., September 2012. Springer LNCS, Springer LNCS. 8, 62
- [23] Shahram Ebadollahi, Anni R. Coden, Michael A. Tanenblatt, Shih-Fu Chang, Tanveer Syeda-Mahmood, and Arnon Amir. Concept-based electronic health records: opportunities and challenges. In *Proceedings of the 14th annual ACM international conference on Multimedia*, MULTIMEDIA 06, pages 997–1006, New York, NY, USA, 2006. ACM. 8
- [24] Council of Europe. Protection of medical data - recommendation no r (97) 5, 1997. 8
- [25] U.S. Department of Health & Human Services. Health insurance portability and accountability act, 1996. 8, 9
- [26] C. Pereira, C. Oliveira, C. Vilaça, and A. Ferreira. Protection of clinical data - comparison of european with american legislation and respective technological applicability. In *HEALTHINF 11*, pages 567–570, 2011. 9
- [27] Republica Portuguesa. Lei acesso aos documentos da administração 46/2007, 2007. 9
- [28] NHS choices. How do i access my medical records (health records)?, 15/09/2010 2012. 9
- [29] Santos-Pereira C. Antunes L., Cruz-Correia R. and Ferreira A. One way to patient empowerment - a proposal for an authorization model. In *In proceedings of the HealthInf 2012 - International Conference on Health Informatics*, pages 249–255, 2012. 9
- [30] Kroll Fraud Solutions. Healthcare information and management systems society (himss) analytics report: Security of patient data. Technical report, Kroll Fraud Solutions, 2008. 9

- [31] Watts J., Huiming Yu, and Xiaohong Yuan. Case study: Using smart cards with pki to implement data access control for health information systems. *IEEE Southeastcon 2010: Energizing Our Future*, pages 163–167, 2010. 9
- [32] Randy Baden, Adam Bender, Neil Spring, Bobby Bhattacharjee, and Daniel Starin. Persona: an online social network with user-defined privacy. *SIGCOMM Comput. Commun. Rev.*, 39:135–146, August 2009. 13
- [33] Sebastian Clauss and Marit Kohntopp. Identity management and its support of multilateral security. *Computer Networks*, 37(2):205 – 219, 2001. 13, 15
- [34] Crc For Enterprise, Audun Josang, and Simon Pope. User centric identity management. In *in Asia Pacific Information Technology Security Conference, AusCERT2005, Australia*, pages 77–89, 2005. 15
- [35] Kim Cameron. The laws of identity. microsoft whitepaper, may 2005. <http://tinyurl.com/kimcameronlaw> verified on 03/10/2012. 15
- [36] G.J. Ahn and J. Lam. Managing privacy preferences for federated identity management. In *Proceedings of the 2005 workshop on Digital identity management*, pages 28–36. ACM, 2005. 17
- [37] Scott Cantor. Shibboleth architecture, protocols and profiles. <http://tinyurl.com/Shibboleth123> Verified on 13/10/2012, September 2005. 18
- [38] K. Cameron and M. B. Jones. Design rationale behind the identity metasystem architecture; microsoft corporation. <http://tinyurl.com/infocards> Verified on 13/10/2012, September 2006. 21
- [39] Microsoft Research unit. U-prove. <http://tinyurl.com/uprovemic> Verified on 14/01/2013., January 2013. 21
- [40] Christian Paquin. U-prove technology overview v1.1 (revision 2), april 2013. <http://tinyurl.com/uprovesep> verified on 03/04/2013. 21
- [41] Abhilasha Bhargav-Spantzel, Jan Camenisch, Thomas Gross, and Dieter Sommer. User centricty: A taxonomy and open issues. *J. Comput. Secur.*, 15(5):493–527, October 2007. 21
- [42] Eclipse Foundation. The project higgins: Personal data service. <http://www.eclipse.org/higgins/> Verified on 14/10/2011, April 2011. 21

- [43] G. Inman and D. Chadwick. A privacy preserving attribute aggregation model for federated identity managements systems. *Serbian Publication InfoReview joins UPENET, the Network of CEPIS Societies Journals and Magazines*, page 21, 2010. 23, 24
- [44] D.W. Chadwick. Authorisation using attributes from multiple authorities. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2006. WETICE 06. 15th IEEE International Workshops on*, pages 326–331. IEEE, 2006. 23
- [45] Apurva Mohan and Douglas M. Blough. Attributetrust a framework for evaluating trust in aggregated attributes via a reputation system. In *Proceedings of the 2008 Sixth Annual Conference on Privacy, Security and Trust*, PST 08, pages 201–212, Washington, DC, USA, 2008. IEEE Computer Society. 24
- [46] Jill Gemmill, John-Paul Robinson, Tom Scavo, and Purushotham Bangalore. Cross-domain authorization for federated virtual organizations using the myvocs collaboration environment. *Concurr. Comput. : Pract. Exper.*, 21:509–532, March 2009. 24
- [47] David W. Chadwick and George Inman. Attribute aggregation in federated identity management. *IEEE Computer*, 42(5):33–40, 2009. 25
- [48] J. Kohl and C. Neuman. The kerberos network authentication service (v5), 1993. 25
- [49] Oasis. Oasis security services (saml) tc. <http://tinyurl.com/SAMLtse> Verified on 14/10/2012. 26
- [50] Oasis. Oasis extensible access control markup language (xacml) tc. <http://tinyurl.com/xacmlTese> Verified on 14/10/2012. 27
- [51] Oasis. Web services federation language (ws-federation). <http://tinyurl.com/wsFederation> Verified on 14/10/2012., May 2009. 27
- [52] TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU. Recommendation itu-t x.1251. <http://tinyurl.com/itustandard> Verified on 14/10/2012., May 2009. 28
- [53] Open Identity Exchange. Oix:open identity exchange. <http://openidentityexchange.org/> Verified on 14/10/2012. 28

- [54] Bruce Schneier. *Applied cryptography (2nd ed.): protocols, algorithms, and source code in C*. John Wiley & Sons, Inc., New York, NY, USA, 1995. 29
- [55] Luís Maia and Manuel Eduardo Correia. Java jca/jce programming in android with sd smart cards. In *7th Conferencía Ibérica de Sistemas y Tecnologías de Informaci3n (CISTI 2012)*, Madrid/ Spain, 2012. 33, 62
- [56] G&D Secure Flash Solutions. Mobile security card. <http://tinyurl.com/msctese> Verified on 14/10/2012., 2010. 33
- [57] P. Saint-André, K. Smith, and R. Tron, con. *XMPP: the definitive guide*. Definitive Guide Series. O'Reilly, 2009. 35
- [58] Peter Saint-Andre Ian Paterson. Xep-0206: Xmpp over bosh. <http://bit.ly/xep0206> Verified on 14/10/2012, July 2010. 36, 41
- [59] M. Kalin. *Java Web Services: Up and Running*. Oreilly and Associate Series. O'Reilly Media, Incorporated, 2009. 36, 37
- [60] Hsiang-Cheh Huang, Feng-Cheng Chang, and Wai-Chi Fang. Reversible data hiding with histogram-based difference expansion for qr code applications. *IEEE Trans. Consumer Electronics*, 57(2):779–787, 2011. 38
- [61] S. Crocker D. Eastlake, J. Schiller. Randomness recommendations for security. <https://ietf.org/rfc/rfc4086.txt> Verified on 14/10/2012, June 2005. 41
- [62] Yuan Cao and Lin Yang. Gisl: a generalized identity specification language based on xml schema. In *Proceedings of the 7th ACM workshop on Digital identity management, DIM 11*, pages 3–12, New York, NY, USA, 2011. ACM. 42
- [63] Mohamad Badra, Ahmed Serhrouchni, and Thomas Guillet. Random values, nonce and challenges: Semantic meaning versus opaque and strings of data. In *VTC Fall*. IEEE, 2009. 43, 54
- [64] Yewsiang Poong, Khaliq-Ul Zaman, and Mohammad Talha. E-commerce today and tomorrow: a truly generalized and active framework for the definition of electronic commerce. In *Proceedings of the 8th international conference on Electronic commerce: The new e-commerce: innovations for conquering current barriers, obstacles and limitations to conducting successful business on the internet, ICEC '06*, pages 553–557, New York, NY, USA, 2006. ACM. 55
- [65] CEN/ISO EN 13606-4. Health informatics - electronic health record communication - security, 2009. 57