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

Alexandre B. Augusto, Manuel E. Correia
aaugusto@dcc.fc.up.pt, mcc@dcc.fc.up.pt

Center for Research in Advanced Computing Systems (CRACS-INESC TEC);
Department of Computer Science, Faculty of Science, University of Porto;
Rua do Campo Alegre, 1021/1055
Porto, 4169-007
Portugal

ABSTRACT

The massive growth of the Internet and their services is currently being sustained by the mercantilization of users’ identity private data. Traditionally services on the web require the user to disclose many unnecessary sensitive identity attributes like bankcards, geographic position or even personal health records in order to provide a service. In essence the services are presented as free and constitute a means by which the user is mercantilized, often without realizing the real value of its own data for the market.

In this chapter we describe OFELIA (Open Federated Environment for Leveraging of Identity and Authorization), a digital identity architecture, designed from the ground up to be user centric. OFELIA is an identity/authorization versatile infrastructure that does not depend upon the massive aggregation of users identity attributes to offer a highly versatile set of identity services but relies instead on having those attributes distributed among and protected by several otherwise unrelated Attribute Authorities. Only the end user, with his smartphone, knows how to aggregate these scattered Attribute Authorities identity attributes back into some useful identifiable and authenticated entity identity that can then be used by Internet services in a secure and interoperable way.

Keywords: User centricity, Mobile Identity, XMPP, OpenID Connect, Attribute aggregation, Access control, mobile device and Identity management.
1 INTRODUCTION

The explosive growth of the Internet is accelerating the migration of essential real world and monetary infrastructures to the virtual world, with digital identity playing a central catalyzing role for this societal transformative process. Arguably the digital world is radically different from the real world, but there are some essential concepts that are readily transposed. Very much like in the physical world, in the Internet we have people interacting with other people and non-human computerised entities, under highly diverse situations. In the real world people behave rather differently when they are at work, in the grocery store or at the gym, where they assume different roles in the face of different contextual situations. This essential social ability to contextually change the way we relate with others is what must be transposed from the physical world to the Internet every time we try to dematerialise societal real world processes to the virtual world.

A digital identity can thus be readily defined as the “set of characteristics that uniquely describes a digital subject or entity and its relations with other entities or digital subjects in a virtual world”. A digital subject, or entity, is therefore something, not necessarily human, that makes a request in order to access a particular resource (a web page, an item from a database…) and is composed by a set of personal data attributes that in some sense characterizes that person or entity, usually referred to as a “user”. The subset of personal data attributes needed for a specific role (or “user”) depends on the situation and context at hand and is usually referred to as an identity persona (Baden, Bender, Spring, Bhattacharjee, & Starin, 2009). The association between an identity persona and a user is done by the means of an authentication process that can also be conducted by an Identity Management System (IdMS) (Hai-Binh & Bouzefrane, 2008).

Digital identity management systems, like their real world analogues, are essential in ensuring that a network infrastructure is capable to scale and meet the basic interoperable expectations and functionalities concerning security, privacy and reliability that emerge every time there is a need to plan and deploy a well engineered Internet service.

1.1 Digital identity management

Digital identity is maintained by identity management systems (IdMS). These are composed by governing organization policies, economic model, business processes and technologies that implement and manage the personal identity users attributes that are needed to establish and manage access rights to organizational digital assets (David W. Chadwick, 2009). Moreover Identity management systems are also responsible for the digital identity lifecycle management within organizations, as they provide the flexible and scalable means by which it is possible to validate and exchange the digital personal data attributes that one needs in order to establish and promote
interoperability among different systems, in accordance with some set of pre-established organizational security and legal policies. According to Kim Cameron, every useful IdMS should follow the seven “Laws of Identity” (Cameron, 2005) that can be observed on Table 1.

Identity management systems are employed by identity providers (IdP) (Clauß & Köhntopp, 2001) to manage digital identity within an organization, group of

Table 1: The seven “laws of identity”

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

organizations or even the whole Internet. Depending on the scale, their interim structure and the social and/or financial benefits accrued by their deployment; IdPs can be further classified as:

- Traditional (digital silo): Where each service domain deploys its own IdP, thus forcing the user to create multiple independent accounts in order to access different services.

- Centralized: Bringing the concept of single sign-on (David, 2006), later extended with the usage of information cards (Cameron & Jones, 2007) in order to establish a way to dismiss the typical login/password scenario. In this model only one centralized IdP is needed to provides the necessary user credentials to grant authentication for different domains (OpenID (Sakimura, Bradley, Jones, Medeiros, & Jay, 2012)).
• Federated: Where there exists a pre-negotiated circle of trust between the participating administrative domains whose IdPs can then grant access to any one of the service domain that falls within the federation authority as a whole. Authentication within the federation is achieved by presenting a valid identifier emitted and authenticated by any IdP that falls in the circle of trust, creating an asymmetric trust relationships among the members of the federation (Orawiwattanakul, Yamaji, Nakamura, Kataoka, & Sonehara, 2010).

• User-centric: Its main objective is empowering users by returning the identity data control back to the user, the legitimate owner (Higgins (Hai-Binh & Bouzefrane, 2008)) (Bhargav-Spantzel, Camenisch, Gross, & Sommer, 2007).

1.2 Identity providers evolution

The old traditional IdP silo model is still the most commonly deployed type of IdP currently in use on the Internet. It requires the user to manage a set of different credentials for different services, which leads to well known security issues (Hovav & Berger, 2009) and intractable interoperability problems. There is however one positive property that emerges from the widespread usage of silo based IdPs, which is the real and effective fragmentation of the user digital identity attributes among different unrelated Identity providers. This is positive from the users’ privacy point of view, since under the silo model no single system can own a complete full set of the user’s identity attributes. In other words, the users’ personal data is naturally decentralized and this helps to protect and improve upon the users’ privacy. However the silo model in its current form does not scale in the Internet. Not only it makes it very difficult to build effective interoperability among otherwise unrelated systems but it also constitute an obstacle to the implementation of secure single-sign-on (David, 2006) mechanisms in a standardized way.

More recently the interest on housing more comprehensive sets of users identity attributes under the same roof, has been increasing dramatically due to the discovery of their highly strategic and commercial value for the Internet market (Schwartz, 2004). Massive centralized identity providers have started to flourish on the Internet. 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 strategic purposes is to create enormous monopolized centralized databases of their users identity attributes, as they allow them to produce highly accurate user profiles that they can then monetize very efficiently for marketing and further lock-in purposes (Schwartz, 2004). These global companies harvest and aggregate personal data in such a massive scale that, lest it is put under some kind of restrain, it will very soon represent a major global threat to personal security and privacy the like of which the world has never seen.

This competition over digital identity led to the emergence of new standardized identity management protocols like OpenID (Sakimura et al., 2012), for interoperable users authentication and identity management and OAuth (Hammer-Lahav., 2012), for authorization management. These protocols opened the way for new network
infrastructures that help cater the need web applications have for data interoperability. These protocols provide the necessary means for centralized identity providers to operate in standardised and interoperable ways and are employed as standard mechanisms upon which it becomes feasible to build interoperable single sign-on systems and attribute sharing based on the concept of valet keys in an effective way. More recently a new open standard, OpenID Connect (Sakimura et al., 2012), has emerged as a single solution for combining both authentication and authorization within a single standardised infrastructure protocol.

In the middle of this dispute lies the user, most of the time unaware that it is his valuable identity and privacy that is paying for the set of otherwise “free” essential internet services (Facebook, Google, etc.) that he is using on a daily basis, most of the time without his explicit consent or control (Mont, Pearson, & Bramhall, 2003). The user is rarely provided with the opportunity and means to negotiate the real value of something as intrinsically vital as his identity. However this privacy abuse is not the only problem that results from unrestrained data aggregation. If a massive centralized identity provider suffers an attack, millions of highly detailed personal attributes can be immediately compromised with highly severe consequences for the users. All these issues constitute the main motivation behind our proposal for a real time fully distributed mobile based user-centric aggregation IdMS for authentication and authorization.

1.3 User Centric as a solution for users privacy

More recently, the general tendency has been to concentrate development efforts on identity management models. These are being structured around user-centric concepts, totally in concert with a more interventive and democratic digital society 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 show 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 the users basic right for privacy.

There are also some types of sensitive 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 highly dynamic personal data like GPS, heart beating and etc. For these reasons we believe that all users’ private information should be kept as much as close as possible to their owners' primary source, under the user’s direct control. For example if the data is the user’s current GPS position, this data should not be shared with a third party node like the Fire Eagle (Inc, 2007) in order to deploy the information to a relying party (a service that consumes the users’ identity in order to obtain a set of users’ identity attributes). In order words, the relying party should be able to directly request the GPS coordinate from the users original data source (in this case his GPS device) and not the last time that this individual or application remembered to update its value from an otherwise stale source of information (in this
case, a positioning identity attribute stored in a traditional IdPs). Highly dynamic personal data should therefore be securely disclosed on demand by the owner’s original data source directly to the requester relying party at the users discretion.

Access to the owners data primary source is managed by the user by engaging the help of a personal Authoritative Authority (Paci, Shang, Jr., Fernando, & Bertino, 2009) (AA). These AAs are entities in the network that disclose personal data to other relying applications at the users discretion after their explicit consent, which must be informed in the sense that it must be based on a reliable trusted identity for the original requester at the relying party. These authorization consents should also be limited in time and be easily revoked. These constitute basic essential assumptions for a well designed user-centric attribute aggregating identity management system.

1.4 Attribute aggregation model to decentralize data

As previously discussed, silo identity management models suffer from serious security issues common to all centralized systems. If the identity attribute storage model is fully centralized it could potentially become a victim of targeted attacks that could compromise the entire user’s digital identity. It is therefore an intrinsic design directive for our aggregation model to have the user’s digital identity split and distributed between different data primary sources. There are many user centric attribute aggregation models to choose from. The most relevant in the literature are:

- Identity relay (Inman & Chadwick, 2010): The Relying Party (RP) trusts a single master federated IdP, that is responsible to request and relay all attributes to the RP, who is then responsible for their aggregation.

- RP mediated attribute aggregation (Sakimura et al., 2012): The RP redirects the user-agent to each IdP thus obliging the users to a high level of interaction and making the user-agent responsible for attribute aggregation.

- Client mediated assertion (Inman & Chadwick, 2010): Based on an intelligent user agent that guide the user to the different IdPs, obliging the users a high level of interaction, the user agent is responsible for the attribute aggregation and the delivery to the Relying Party.

- Identity Federation model (David, 2006): After user authentication, a secret is generated and shared between all federated IdPs by the user-agent thus allowing the RP to request the needed attributes from all the federated IdPs.

- Identity proxying/chaining (Gemmill, Robinson, Scavo, & Bangalore, 2009): The RP fully trusts in a single master federated IdP that is responsible to request and aggregate all requested attributes.

- Linking Service (D. W. Chadwick & Inman, 2009): In this model only the user knows about all his IdPs, a service called linking service is responsible to hold
minimal information that allows RPs to obtain their queries from the other IdPs via the linking service. After user authentication the IdP offers the possibility of attribute aggregation and if the user accepts it, the information to access the linking service is shared with the RP. The aggregation of attributes can be done by the linking service itself or at the RP.

Our proposal, OFELIA, is based on a user identity attribute aggregation infrastructure that falls within the “Linking service” category (Augusto & Correia, 2012). In OFELIA attribute aggregation is realised by the means of a secure mobile authentication and authorization broker, running on smartphones, where users exercise discretionary asynchronous control over access requests to their personal identity attributes. These are distributed and located among several different attribute storage network nodes that have been previously established as Authorization Authorities (AA). The smartphone thus provides the user with the means to also exercise aggregated management control of his AAs, by conditionally making them accessible to web applications (Relying parties) on behalf of its authorised trusted users and during a predetermined, but revocable, well defined period of time.

The adoption of smartphones as a user-centric management platform is highly appropriate because these devices are nowadays ubiquitous, have more than adequate processing power, provide Internet connectivity and follows his owner everywhere, thus providing a practical solution for the users Internet reachability challenge (Barkhuus & Polichar, 2011). The use of smartphones for identity management is currently also recognized as essential for enhancing security and privacy (Adi, Al-Qayedi, Zarooni, & Mabrouk, 2004; Paci et al., 2009; Zhikui, 2007) and has been proved to play a crucial role on the more flexible user-centric models (Augusto & Correia, 2013).

1.5 The OFELIA proposal

Our overall goal is to define and specify a fully decentralized privacy and user-centric infrastructure for identity management based on the distributed aggregation of users private data and protected by a set of personal AAs. In OFELIA the user personal smartphone acts as the Linking Service where the user directly manages attribute aggregation and access authorization.

We also intend to deploy the user smartphone as a secure authorization broker where attribute aggregation is achieved by securely enrolling the users AAs and their respective managed identity attributes into the users smartphone. To implement and deploy OFELIA we relied on already proven and standardised protocols/infrastructures like:

- Extensible Messaging and Presence Protocol (XMPP) and Restful web services: to authenticate, validate and establish the communication between the intervening network nodes.

- OpenID: employed as an authenticator and provider of the necessary bootstrapping information about the user.
Quick response code: to quickly exchange digital information with the smartphones in order to simplify user experience.

Valet key based protocols: to create the necessary means for managing and conveying conditional, but revocable access authorizations to Relying Parties.

Public key infrastructure: to manage trust among the different participating network nodes.

MicroSD mobile security cards: a smartphone mobile smartcard to better secure the linking service privately held keys and provide second level token based authentication capabilities to the Identity Attribute Aggregation service.

We are currently actively engaged in developing and deploying the following four different components: (1) one application programming interface (API), to allow for a faster and simpler deployment of third party Relying Parties and Service providers into OFELIA; (2) other API for the enrollment of third party Attribute Authorities, thus helping to diversify the universe of available identity attributes; (3) an implementation for an Identity Broker (Augusto & Correia, 2012), another component of our linking service aggregation model; (4) and an android application, implementing the secure mobile authorization broker to enroll and aggregate AAs and authorize, manage and revoke access to users identity attributes that are being managed and secured by the enrolled AAs.

The rest of this chapter is organised as follows. In section 2, we review the proposed architecture, describing each technology and their functionality in the architecture components that are described as well. In section 3 we describe in detail our protocol responsible to establish the connection between the architecture components and describe a usage case scenario, which can be quite useful to help to better understand the different components interaction. In section 4 we described what has already been accomplished and present some preliminary conclusions for the work we have already developed thus far within our project OFELIA (Open Federated Environments Leveraging Identity and Authorization). We finalize with a brief outline of our future development and discuss some implementation notes about libraries and software that we have used during the course of the implementation of OFELIA.

2 ARCHITECTURE

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 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. In what follows we provide a more detailed description of the functional role played by each one of these parts.

![Figure 1: OFELIA architecture](image)

2.1 The Relying Party/Service Provider (RP/SP)

The RP/SP is a web application that requires user’s 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 user’s 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 digital identity XML semantic specifications. The RP/SP should also provide safe caching of authorization tokens while their validity assertions holds true.

2.2 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 user’s personal data aggregation and announces to the IdB 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 IdBs 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 AA’s 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.
2.3 The Identity Broker

The Identity Broker (IdB) exists to cater for privacy enhanced contexts where the RP/SP cannot be fully trusted and to prevent the more popular AAs to directly track users while they navigate through web applications that use as part of their digital identity and personal data infrastructure. Moreover for privacy and security reasons the IdB must also not know the content of the personal data it is relaying. This is accomplished by having the RP/SP and AA to negotiate session keys and then encrypt all personal data that is being relayed by the IdB.

The proposed architecture aims for a trust balance where the RP/SP does not have to know about the aggregation of AAs and the IdB 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 IdB to be in the possession of a public key pair whose legitimacy can be attested by a valid PKI X509/PGP certificate with the IdB identity.

2.4 The smart-phone as an Authorization Broker

In OFELIA we are employing android smart phones 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 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 applications 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 terminates 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 section 3.

2.5 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 (Cao & Yang, 2011).

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, which we employ throughout our implementation to keep and promote interoperability for data exchange within OFELIA. The Figure 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.

![Figure 2: OFELIA data exchange Schema](image)
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 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** (Badra, Guillet, & Serhrouchni, 2009); 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 and among others 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 **OfeiliaDataExchange** 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.

### 2.6 XMPP: Extensible Messaging and Presence Protocol

XMPP is an open technology for real-time communication that uses the eXtensible Markup Language (XML) as a base format for exchanging information formatted as XML documents. These documents are sent from one entity to another (Saint-Andre, Smith, & Tronon, 2009) by using an appropriate application level transport protocol according to network availability. XMPP servers provide a very flexible set of standard services that can be used by many different types of applications like network games, chat system, etc.

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 (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, a XMPP messaging infrastructure proves to be an almost ideal communication asset for our needs because of its core services, namely:

i. Almost real time messaging. Essential to maintain accurate and updated the dynamic data types that are being maintained and exchange between the Identity Broker, the Authorization Authorities and the Authorization Broker (smartphone).

ii. Authentication by digital certificates. Guarantees a high level of trust and non-repudiation between architecture nodes and users.

iii. Ability to efficiently operate over HTTP by the means of the BOSH (Bidirectional-streams Over Synchronous HTTP) protocol (Paterson & Saint-Andre, 2007), where two non directly addressable devices located on private closed intranets, with minimal Internet access, can locate each other over the Internet and then directly exchange messages in a reliable and safe way.

iv. Its capacity to store and forward messages in case any of the nodes becomes offline. This proved to be a very strong and convenient asset to have for asynchronous communications with mobile devices. It is important to have in mind that cellular phones are often located in areas with bad data coverage, which results in severe communication problems that have to be dealt in an asynchronous manner.

v. Its scalability, to avoid bottleneck problems and the fact that it is a mature, fully supported and approved Internet standard that is widely deployed and is currently a very important part of the communication operations and infrastructure of large distinct companies like: Google, Facebook, Blizzard, Steam and among others.
2.7 OpenID

The OpenID provides a decentralized protocol for user authentication. It is deployed as part of Identity Managers that allows a user to sign into distinct domains with a single OpenID account (single sign on) and at the same time let the user control what of his identity attributes will be disclosed in order to identify and authenticate himself into the domain that is acting as an OpenID consumer.

In order for a user to authenticate into a domain with OpenID, he needs to be redirected to his OpenID provider where he is asked to authenticate (usually via a login/password method) and then authorise the identity attribute exchange requested by the domain. If this proves to be successful the user is then redirected to the originating requesting domain and granted access. In order to standardise and define appropriate semantics for a minimum useful set of user attributes that could be universally recognised by all RPs, the full set of standardised and widely recognised identity attributes for OpenID is unfortunately substantially small. This decreases the usefulness of the protocol and has so far limited its deployment almost exclusively to the authentication domain.

Recently the OpenID foundation started to work in a new protocol named OpenID Connect that aims to unify authentication and authorization in a single service protocol. This unification will create the right means for data access authorization and it will prove to be a firm step towards solving the issues resulting from a too limited set of widely recognized identity attributes.

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 user's 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.

2.8 Quick Response code

A 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 (Hsiang-Cheh, Feng-cheng, & Wai-Chi, 2011).
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 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.

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 our authorization broker (smartphone) and the other nodes of the OFELIA infrastructure.

### 2.9 Valet key based protocol

Nowadays many common authorization protocols like Kerberos and OAuth 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 valet key for data access in the sense that any entity that possesses the key has temporary and restricted access to the protected resource. One of the most common scenarios 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 anytime by the user that owns or manages the data.

For security reasons the authorisation token must be very hard to falsify. In OFELIA it takes the form of a base64 encoded XML excerpt, containing elements for a large pseudo-random number (Eastlake & Schiller, 2005) 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 user’s smartphone private key. The resulting XML document is then encoded into a base64 string, which then constitutes a well formed OFELIA token.
These valet key 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 like login/passwords. In OFELIA these authorisation tokens are issued by the authority broker (user’s smartphone) and are only 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.

2.10 Public key 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 (user’s 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 are 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.

2.11 MicroSD mobile security card

Due to its potential economical factor (Barkhuus & Polichar, 2011), the hunger for mobile devices that can act as an authentication/authorization node are daily increasing. Mobile operators like Orange 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 service 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.
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 (Maia & Correia, 2012) and its protocol interface assure 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.

Nowadays almost all smartphones accept the microSD card in order to expand its storage capacity. 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 with JavaCard OS. This card has a special place in our architecture since its responsible for guaranteeing a strong user authentication and trustworthy protection of data. Otherwise we would to rely on a regular file based 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 posses a copy of the keystore file in order to proceed with the attack.

3 Enrollment processes and usage case scenario

In OFELIA 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 aggregations participating Attribute Authorities that manage the user’s 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 user’s 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 OFELIA architecture.

3.1 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 painlessly 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 subsection 2.) to properly engage with the other infrastructure participants.

OFELIA provides AAs with an easy and secure method to help the user link his smartphone to the AAs accounts that make up the user’s 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 user’s 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 the smartphone with a particular AA the user only has to start an authenticated web session with the particular 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. Figure 3 exemplifies the AA enrolment process providing a more technically detailed description of the whole process.
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 user’s public key, all compiled and encoded as a QR code.

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

3.2 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 user’s 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 logins/authenticate into the IdB with OpenID Connect account, which provides the IdB with the XMPP identity (jabber address) and the public key of the user’s 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.

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 also use the OFELIA App to revoke previously issued and still valid authorizations tokens.

**Figure 4: 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 all the attribute names and respective types that are being aggregated by the users smartphone.

7. The IdB confirms the registration to the user’s smartphone and this concludes the mobile phone IdB enrollment process.

3.3 Relying service 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 user’s identity attributes and allows the RP to enroll with IdB as a user’s 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 user’s data for that RP. This triggers an authorization request from the IdB to the user’s’ smartphone that must be acted upon by the user and leads to the issuing of authorization tokens by the smart phone.

On the user’s approval, the OFELIA 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. This scenario is exemplified in Figure 5. Now the RP can request attributes from IdB while the authorization given by the user remains valid.
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 data enciphered with 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 (Badra et al., 2009) 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).

Figure 5: Relying Party enrollment flow
7. At the user’s 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.

3.4 Usage Case scenario

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 let's 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 user’s 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 sent 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 CONCLUSIONS

With the proposed infrastructure 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 temporal constraints by the arbitrage of the user’s smartphone.

OFELIA 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 IdB precisely to prevent the massive disclosure of user data lest the IdB be compromised. In OFELIA if an attacker compromises the IdB he will not have disclosed the user’s 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. This opens a whole new range of opportunities and possibilities due the ability to allow data be processed as requested. In other words every time a relying party requests an identity attribute, this data value is processed in real time, becoming an essential feature for dynamic attributes that for its own nature is volatile.

4.1 Future Work

We are currently extending OFELIA with smartphone to smartphone communication mechanisms parameterized by QR codes to cater for side channel authorization requests in the case where some OFELIA user, enrolled in a relying party and acting as some predefined role wants to directly ask to some other user, permission to access some of his OFELIA managed identity attributes.

Interoperability between different identity management systems is the key for usability. Therefore we intend to research and develop a novel XML based digital Identity model to improve upon the main ideas present on other semantic models for user-centric identity and base it on SAML (Saklikar & Saha, 2007), metadata identity semantics and other alternative distributed digital identity semantic models (Cao & Yang, 2011).

4.2 Development notes

As already mentioned, to implement OFELIA architecture we relied in some libraries and software. In this subsection we exposed the libraries, their versions and if possible their respective download links:

OpenID consumer library:
Openid4java
Version: 0.9.5.593
Download link: http://openid4java.googlecode.com/files/openid4java-full-0.9.5.593.tar.gz

XMPP BOSH client connector:

Ignite realtime SMACK API
Revision: 12894
Svn link: http://svn.igniterealtime.org/svn/repos/smack/branches/bosh/

XMPP test server:

Ignite realtime Openfire
Version: 3.6.4
Linux download link: http://www.igniterealtime.org/downloads/download-landing.jsp?
  file=openfire/openfire_3_6_4.tar.gz

Android API 10 for android 2.3.3
Android SDK
Version: 20.0.3
Linux download link: http://dl.google.com/android/android-sdk_r20.0.3-linux.tgz

Acknowledgments

This work is funded by the ERDF through the Programme COMPETE and by the Portuguese Government through FCT - Foundation for Science and Technology, 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.

REFERENCES


**Key terms:** User centricity, Mobile Identity, XMPP, OpenID Connect, Attribute aggregation, Access control, mobile device and Identity management.