Design a platform for providing home-based services concerning: care, safety, comfort and fun

MASTER THESIS

Diego Vettorel
diegovettorel@gmail.com | +31 619988253

Public version

Master of Science in Business Information Technology
School of Management and Governance
University of Twente
Enschede, The Netherlands
17 July 2015

Graduation Committee:
Dr. Ir. E.J.A. Folmer, University of Twente
Dr. Ir. Marten J. van Sinderen, University of Twente
Drs Paul Francissen, Envolve owner, thesis and internship supervisor and HuisKluis concept's creator

HuisKluis concept's creators
Paul Francissen - paul.francissen@envolve.nl - +31610714772
Rob Van Dort - rob.van.dort@gmail.com - +31629525229
John Van Echtelt - john@van-echtelt.nl - +31610850096

UNIVERSITY OF TWENTE. HUISKLUIS envolve management en advies

Preface

This public version is a condensed version of the complete version that will be available for the public the 17/07/2017. This version contains only part of the contents of the full version.

Everything begun in October 2014. Rob Van Dort explained to me, during the University of Twente Hackathon, the idea and concept of HuisKluis (the foundation that provided the business case for this thesis). I was fascinated by the idea and we decided to meet with Paul Francissen to discuss it. When we met, Paul and Rob told me their idea to make a platform to import, see, use and share data of end-users about their houses. I was enthusiastic about the idea and we decided that it would be a great project for my master thesis. This project took the form of an internship of six months and was financed by the Envolve company together with Foundation HuisKluis. Envolve was one of the initiator of the HuisKluis foundation. I worked with them for six months to make this thesis that summarizes all their ideas and concepts and includes my analysis as well.

Paul, Rob and John Van Echtelt had this beautiful idea and I personally think that it is a great idea that will become a successful platform. During my internship they were always willing to give me help, information, support and feedback.

This thesis is the summary of all the ideas, concepts, use cases, interface design and end-user's analysis that they had, and adds my analysis of and designs for those concepts. Every time that "the platform" is mentioned, I am referring to the HuisKluis platform. This thesis contains all the ideas that refer to the HuisKluis concept. Therefore the thesis will remain confidential and not public for 24 months from the date of my graduation.

This thesis was possible just because of the time and effort that all the different participants put in it. I would like to say thank you to my project supervisors Paul, Rob and John. We did a tremendous amount of work together in developing this project. We did a great job guys, this will be a successful project when it will be realized. A special thanks to Envolve for financing this thesis, without which this thesis would not have been possible.

If you are reading this thesis and would like more information about the concept you can contact the following end-users:

Diego Vettorel - diegovettorel@gmail.com - +31619988253 (author of this thesis)
Paul Francissen - paul.francissen@envolve.nl - +31610714772 (HuisKluis concept's creator)
Rob Van Dort - rob.van.dort@gmail.com - +31629525229 (HuisKluis concept's creator)
John Van Echtelt - john@van-echtelt.nl - +31610850096 (HuisKluis concept's creator)

I would like to thank my university supervisors Marten Van Sinderen and Erwin Folmer. They always helped me with a lot of useful feedback to improve the thesis. They were always honest and reasonable in the development of the thesis and their contribution to the project was indispensable.

I would like to thank the University of Twente that offers this master program. I think that the Business Information Technology Master program really helped me in expanding my knowledge of both the business and the IT side. I think that it is a unique combination that I hope will give me a brilliant future in the IT and business environment.

I want to say thank you to all the people that supported me in my Dutch university career. My parents Roberto and Djamila always supported me during all my studies. My flatmates Sigrid, Marga, Vincent, Rene and Danielle were always willing to help me and to support me in any situation. Thanks guys for all the support. Thanks to Michele for the support that you have given me in the last years, that was very important to me. Finally thanks to all my friends that in the past years supported me in everything, sorry if I do not mention everybody here, the list will occupy a lot of pages. Guys, I would not have succeed without you. Thanks a lot.

Enschede, 17 July 2015
Diego Vettorel

Abstract

Currently the information about a house is scattered over different websites and stored on different servers on the internet. House owners, inhabitants and renters perceive this to be a problem. Additionally, they often have problems sharing data with third parties in a simple and secure way, without any external applications using their private data without their permission. Besides end-users, several stakeholders are involved in this problem such as service providers and data providers. These different stakeholders need to have a platform where their needs can be met by cooperating with each other to offer the best service possible.

The problem that is addressed in this thesis is how to create this platform: to decide which services can be implemented in the platform that will use the end-users' data, to choose the software architecture and how to solve the trust issue between end-users and the different parties involved.

This thesis investigates these problems and provides a solution for each of them. Research pointed out which requirements the platform will need to have to fulfill the needs of the stakeholders. These requirements include different levels of authentication for end-users, the possibility to import external data in the platform, the possibility to develop external services that can be integrated in the platform, the possibility to share end-users' data with external websites and the alert function.

Several services that will use the end-users' data are described, such as the alarm sharing service, the photo retrieval and upload service, the company information service and a service that provides information about the appliances in the house. Additionally, there are many types of data sources that can be imported in the platform such as sources that give information about security, energy consumption, flood risks, garbage collection and municipality contact information.

Mockups and a software prototype of the platform are provided to show the different parts of the user interface and how the platform will work. The software architecture of the platform is divided in different parts. The first part is the user interface that includes different ways end-users will be able to access the platform, for instance by using a tablet and mobile browsing. The second part is the platform where the different services and servers are located. These services and servers include the linked data server, the file server, the data storage and retrieval service, the authentication service and the data conversion service. The third part consists of external actors that are connected to the platform. These actors include external websites with which users are able to share their data, the external alert function, the static data sources and the dynamic data sources.

The trust issue that plays a part in making this platform a success is also addressed. End-users do not want to share personal data with companies and other parties without their consent. However, data sharing by end-users is an essential assumption of our project, and therefore do users need to feel safe about sharing their data with the external services of the platform.

This thesis proposes a fourfold solution to this problem. The first solution is to create a foundation that will own the platform. The foundation will have the role of controlling and supervising the development and the governance of the platform. Since a foundation is not able to share its profits with shareholders, end-users will know that their data is not used for commercial purposes. Secondly, a manifesto of the foundation was written that states that the end-users are the owners of their data and that the foundation cannot share it without their consent. The third solution is to provide different levels of authentication for end-users, this will increase the security of the platform. The fourth solution is that the end-user will be able to control every single data source that they will share with external parties. End-users will have total control over the decision with whom they share their data.

Table of contents

List of abbreviations

OSS:	Open Source Software
MM:	MissionMode
ADS:	Advertisements
SA:	Software Architecture
IdM:	Identity Management
JS:	JavaScript
SOA:	Service Oriented Architecture
CSS:	Cascading Style Sheets
API:	Application Programming Interface
GPL:	General Public License
#:	Number
UI	User Interface
RQ	Research question
R:	Requirement

List of tables

List of figures

1. Introduction

In this thesis we will present a complete business case for the design of a platform that provides the end-user with useful information and functions that will permit him to interact with different data connected with the system and his house. Keywords for these services are care, safety, comfort and fun.
In this thesis we will use the word "use case" to refer to a service that, with sources of information, will provide useful functions to end-users. A "use case" is also a service that provides just information to end-users. In this thesis I will use any plural personal and possessive pronouns (we, our) to describe the work of the author of this thesis (Diego Vettorel), Envolve's owner and HuisKluis concept's creator (Paul Francissen) and the HuisKluis concept's creators (Rob Van Dort and John Van Echtelt). In this thesis we will use the word "inhabitant" to refer to a person that live in a house.

In this introductory chapter we present an introduction on the topic, the motivation that lead us to write this thesis and the research objectives and methodology of this thesis.

1.1. Context

These days, open data is becoming more and more important. We are going to enter a new era: the Web 3.0 era. In this era, data will be better connected to the internet (for instance with Linked Data). People will be able to search and use data stored in different servers in an easy and efficient way. Data will be structured and easily accessible.

The problem is that at this moment an easy and widely used platform to collect and access information does not exist. A new way to manage this standard information will be used in the future by all internet end-users. The situation is similar to the early age of social media. Many different companies are trying to make their interfaces and platforms the dominant factor but none prevails over the others.

The technology and the ideas are already on the market. What the different parts involved really need is an organization that will lead the market and that will provide a real standard and a platform for house-related data and services. This to collect all the information in a simple way and to give access to as many end-users as possible. We have put a lot of consideration into the linked-data technology.

A few different companies are trying to solve this problem, but in this thesis we will focus on the collection of data regarding houses and on providing useful services that use this data.

1.2. Problem statement

The main problem that we want to solve is how to make the life of end-users easier by connecting the different sources of data about houses in one interface. At the moment information about houses and their qualities is scattered over different sources. Accessing these sources separately and collecting the data this way is time consuming and bothersome. With our platform this problem will be solved.

In addition, we will solve the problem of sharing of this data with other websites and with other applications. Right now, it is difficult for a company to access data of a customer in a simple way and respecting the privacy of the

customer. With our system this problem will be solved by permitting every end-user to choose with who he will share his data. Different companies will be able to access the end-user data in a simple way.

In this thesis we will provide an analysis of a future platform that will solve these problems. These kind of problems, extended to all the sources of information and not limited to the house field, will be a growing problem in the next few years.

With our research we will contribute to solve this problem also in others fields than the house-related information. Our contribution will help future research in this field by providing a complete business case, a complete SA and a complete software prototype of the platform.

1.3. Motivation

This thesis shows a future way to connect different sources of information linked to a house with the owner of the house or with its inhabitants. This idea is the essence of the project. The provided information can be used by external services and applications that will benefit the inhabitants and the owner of the house.

At the moment a lot of information about houses is available for end-users. The problem is that this information is stored in different places, and scattered over different companies and different servers. Also it is not simple to access this information by end-users since not all companies provide an end-user friendly interface.

This thesis contains the analysis for a platform where end-users can access their house's data with a simple and end-user friendly interface. They will also be able to use different apps, developed by third party companies, that will use their houses data. Doing so enables end-users to share their data with external websites and give them the opportunity to decide which data they want to see in their personalized interface. Additionally, end-users can connect with external services that will use the information from the portal to provide useful functions to them. These functions are not available with the current products and technologies. We aim to make them possible with this platform.

This project may be the future of house data sharing. Using our platform a person will be able to access the data of his house from a single and integrated interface, and can use external applications that access this data. For instance, a person can choose to share his alarm security information with his neighbors and with the police. In this way, if a malicious person enters his house, the neighbors and the police will be alerted immediately.

Another example can be sharing information about different rooms in a house with the firefighter department of the city. In addition, there are many different other uses of the system.

End-users will be able to access to this platform with different devices such as tablets and smart phones.

1.4. Research objectives

The objective of this thesis is to design a platform to provide home-based services concerning care, safety, comfort and fun. This includes a business case of the system, a detailed SA of the future system and a detailed description of some use cases.

This is a very innovative field. For instance, we have researched how to aggregate data about houses and how to use this data to help end-users' lives. We researched how to link this data with external services and websites.

We have created use cases that, when they will be implemented, will help end-users and will provide them useful functions. We have also done case study research on the existing platforms that are similar to this one.

Currently end-users cannot access a simple interface where all their house data is shown a simple way. They also cannot use this data with external applications by deciding which application can access which data resource. Data providers now do not have an integrated and simple system to provide their data to end-users. Additionally, app developers do not have an integrated source of information from end-users. With this platform these problems will be solved and the different parties can start to provide these services.

Another objective of this thesis is to provide a software architecture design to implement the platform. The software architecture needs to satisfy the requirements that we have gained from analysis of the wishes of different parties involved. We analyzed different parts of the architecture, for instance the central platform, the end-user interface, the data import and the physical layer. The architecture supports the import of linked data and of non-linked data. This data will be processed and made accessible for third party companies and services.

More specifically the objectives of the thesis can be divided into four main categories:

1. The business case:
    - Provide a business case of the platform.
    - Provide a functional description of the different functions of the system.
    - An alert function that will permit end-users to receive an alert on the mobile app (or by e-mail or by any other communication system) when an alert is sent from the system.
    - A wallet function where end-users will be able to share their data information in a simple way. End-users will be able to share the information of their wallet with every application or website that they want.
    - Provide different identification methods for different levels of trustiness of the end-users.
    - Provide the description and the analysis of several use cases for the system.
    - Provide different examples of data sources from where the system will take the data and how they will connect to the system.
    - Provide different examples of ways to access the data (web interface, mobile app).
2. The software architecture:
    - Provide a software architecture for the system.
3. The trust framework:
    - Provide the solution of the "trust" issue between this platform and the end-users (end-users may be scared to share their data).
4. The proof of concept:
    - Provide the design of the interface that the final dashboard (where all the different information of the house will be shown to the end-user) will have.
    - Provide the design of the interface of several use cases.

1.5. Research questions

The research questions are:

- **RQ1:** What are the characteristics of a platform where end-users will be able to see, share and analyze information regarding their houses?
- **RQ2:** Which use cases will be implemented that use information about houses and that will provide useful functions to end-users?
- **RQ3:** What will the software architecture for such a system look like?
- **RQ4:** What can be a solution for the trust issue between end-users and the information sharing platform?

1.6. Research methodology

For this thesis a design-science approach was used. First interviews were held with the different stakeholders involved. Interviews with end-users, public administration services and social organizations, service providers and data source providers shared their needs and requirements. At the same time a discussion between the owners of the platform was held to find out our needs and requirements.

From the interviews and discussion a requirement list was compiled. Then a market analysis of the existing solutions that have already been developed was done to find out if any organizations had already developed a platform with these requirement combinations.

Then the business case of the platform was made using all of these requirements and the SA of the platform was designed. At the same time the an analysis of the "current state of technology and research" was made and the information that was gained here was used for the business case creation and for the SA design. Furthermore, at the same time several analyses of external companies to research the possibilities of partnerships were made and the information gained from these analyses were also integrated in the business case and in the SA design.

The next major step was to design the UI of the platform by making some mockups. After the designing the mockups were tested with two different user interface designer experts that given us feedbacks that we integrated in the mockups.

Then a software prototype was made and it was tested with potential end-users of the platform to find out if the software requirements were satisfied. Then the final software prototype was made.

Then the problem of the trust issue was analyzed and solved using all the information that we gained from the previous steps.

Finally the discussion of the results and the conclusion of the thesis was made.

Figure 1 shows this research method process.

Additionally the E3value model is used to research which stakeholders will benefit from the platform for each revenue model. Using the E3 value model gives insight into the question when the different stakeholders will be satisfied and which revenue streams they will take and give.

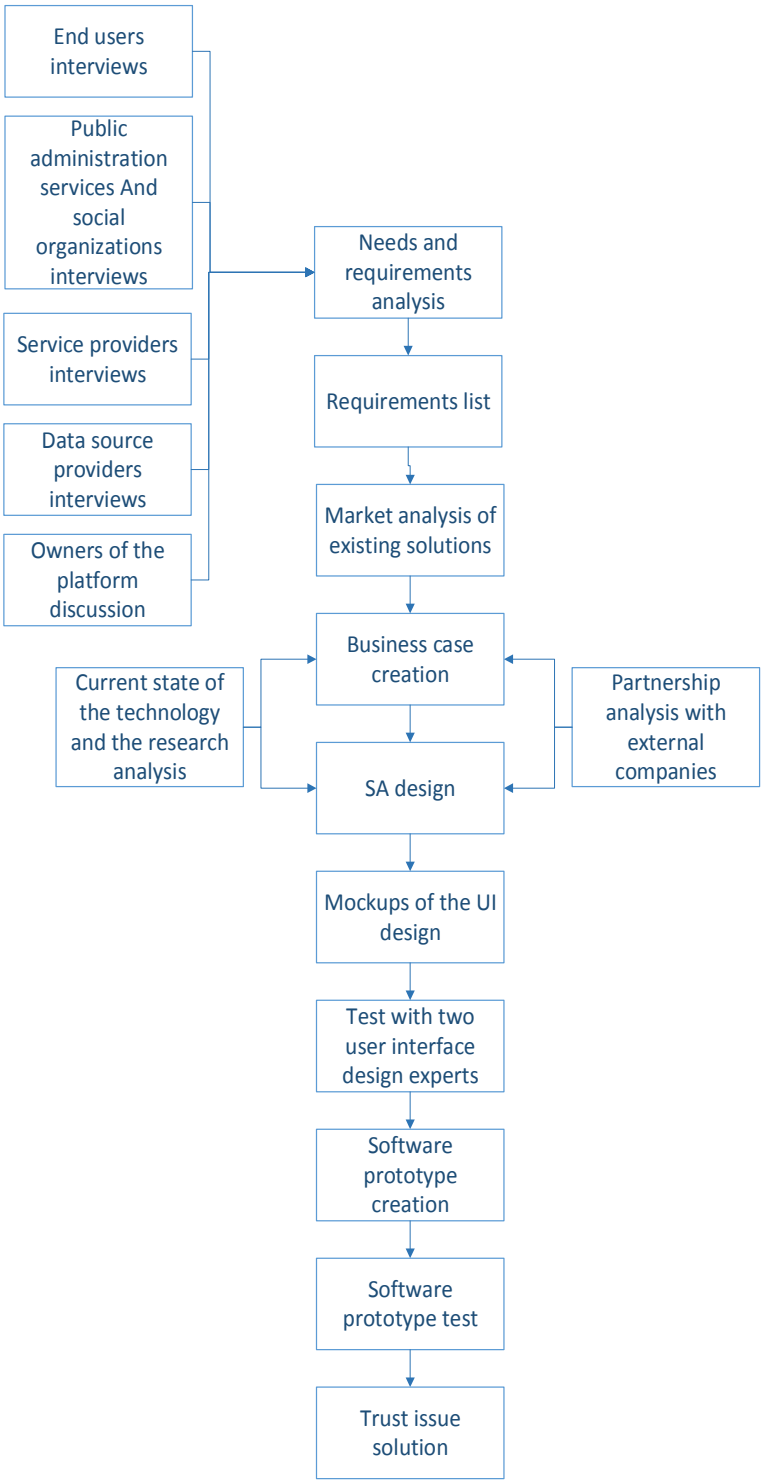

Figure 1: Research framework

1.7. Report structure

In the introduction chapter it is presented a description of the context, the motivation and the methods used for this research.

In chapter two we present an analysis of the problem, the stakeholders and the end-users analysis, the requirements of the future platform and a complete analysis of the current state of the technology.

In chapter three we present the complete description of the business case that we made during our research. A user perspective services analysis was made and there were analyzed some use cases, some data sources and some revenues models. Finally an analysis of the foundation governance model and of the OSS was made.

In chapter four we present the software architecture that the system will have and the process that we followed to make it.

In chapter five we present the description of the mockups of the platform and the analysis of the test that we made of them. The reader can find mockups of the interface that the system will have.

In chapter six the description of the software prototype of the platform is made and some test of it are described.

In chapter seven the trust issue between the platform, the end-users and the external companies is analyzed and there is described why our platform will solve this trust issue.

In chapter eight the discussion of the results is made.

In chapter nine the conclusion is made that includes the review of the research questions, the theoretical and the practical scientific contributions, the research limitation and the recommendations for future work and research.

In the appendixes there are explained some details about the interviews with the end-users that we made, some data sources and some mockups of the interface.

2. Problem analysis and current state of research and technology

In this chapter a problem analysis is provided, along with an analysis of the current state of technology and of research. In the problem analysis part the stakeholders, their needs, their concerns and their requirements are discussed. Then an analysis of the different levels of identity certification for end-users and the end-user requirements analysis is given. Finally, in the current state of research and technology section, existing platforms that provide similar services and functions to this platform are scrutinized. In that section we will also analyze the current state of research and existing technology solutions connected to this platform.

2.1. Problem analysis

2.1.1. Key concept definitions

In this section definitions of the three main concepts are stated. These definitions are the general agreed definitions for the most important concepts contained in this thesis.

2.1.1.1. Integrated platform

An integrated platform is software that integrates different applications and services. Within the boundaries of this thesis "integrated platform" refers to a platform where different data sources, applications and services are integrated.

2.1.1.2. Linked-Data

Linked Data provides a method to publish structured data. The term Linked Data, following Tim Berners Lee (2009), refers to practices for publishing and connecting structured data on the web. These pieces of structured data can be linked together and can be retrieved using the SPARQL language. Currently many sources of Linked Data exist on the internet, which can provide useful data for this platform.

2.1.1.3. Use case

A use case, following Cockburn (2002), is a description of a system's behavior when interacting with the outside world. In this thesis a use case is a service that uses external data sources, new data provided by the end-user, connections with other services and end-user input to provide useful functions to the end-user.

2.1.2. Stakeholders identification, their needs, concerns and requirements

This section has been left out because it is part of the private version of this thesis.

2.1.3. Different levels of identity certification for end-users

The platform needs different levels of identity certification for end-users. Each service and data source will be linked to a different level of end-user certification permitting end-users that have the maximum level of certification to have access to all the available data sources and services. In this thesis the words "authentication" and "certification" are used to express the same concept of having different levels of identity certification for end-users.

To establish these different levels of certification we analyzed the needs of the stakeholders that will provide services and data sources to the platform. This was done by interviewing stakeholders and by analyzing the different services, data sources and functions that they will provide. For each service, data source and function we analyzed the possible level of authentication that the end-users will need to have for each of them. The results are presented in this section.

The different levels of certification for the end-users need to be:

1. **Email confirmation:** this is a low level security measure as anybody can use fake credentials to register to the platform and can use a fake email address.
2. **Address confirmation:** a letter with a security code will be sent to the address of the end-user. The user needs to insert this code into the system to prove that he lives on that address. This level of certification is medium-low because the end-user can use a fake name and surname. Interception of the letter is another threat to this method.
3. **Name, surname and address confirmation with an identity document verification:** end-users will send a copy of their identity document to us. We will verify that the data inserted by the end-user matches the details of the provided identity document. This is a low security certification measure because an identity document image can easily be faked.
4. **Name, surname and address certification from the government**: our system will request certification from the government that proves that the end-user really lives at that address. This is a medium security certification measure because government data is carefully checked. On the other hand, with this security method we can just prove that a particular person lives at a particular address but we cannot prove that the person that registers to the platform uses his real name and surname.
5. **Bank name, surname and address certification:** the credentials of the end-user are checked against his bank account details. This method provides a high certificated identity because every bank end-user is checked personally with his identity document.
6. **DigiD certification:** this certification proves the identity of one person and his address. This provides a high certificated identity because every DigiD identity has the name, the surname and the address of users checked. To have this level of certification we will need to make an arrangement with the Dutch government because private organizations cannot access the DigiD technology.

2.1.4. Software requirements

This section has been left out because it is part of the private version of this thesis.

2.2. Current state of the research and of the technology

2.2.1. Current state of research

In this section current scientific papers relating to our platform are analyzed and discussed. The main catalogue where these papers were found was Sciencedirect (http://www.sciencedirect.com/). Some of the researches presented in this section are not used in the thesis because we realized, during the later analysis, that they were not useful. They are included in this section to show a complete overview of the research the we made.

2.2.1.1. A flexible platform for synchronized measurements, data aggregation and information retrieval

Alexandru Nechifora, Mihaela Albub, Richard Hairc and Vladimir Terzijaa (2015) explained a way to realize a platform that receives measurements from Phasor measurements units (PMUs) and then aggregates and shows their data.

This paper can be helpful for the development of our platform because it explains how to retrieve real time data from measurement units. In our platform we will need to import data from real time sources like thermostats and weather stations. These data then need to be shown to end-users and need to be analyzed in order to provide services and statistics to end-users.

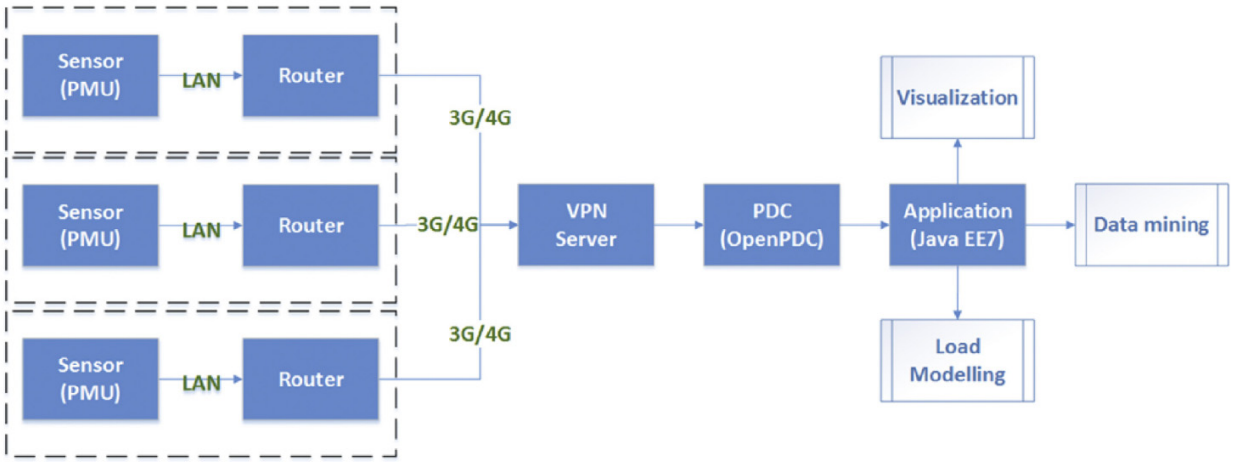

Figure 2: Conceptual view

As shown in Figure 2 above the communication infrastructure proposed in their paper consists of a layer where different sensors are connected to routers through a local LAN network. Then routers are connected to the PDC server using a VPN to increase the security of the connection. The PDC server is then connected to the application layer that is responsible for visualizing and mining the data. Sensors can be connected to the internet with a 3G and 4G wireless connection if there is no possibility to use cables in the position of the sensor.

It is possible to use this infrastructure in the SA of our platform.

2.2.1.2. A virtual sensor system for end-user-generated, real-time environmental data products

Hill et al. (2011) describes a method to retrieve data from sensors. The prototype that they developed receives real-time data from sensors, enhances it and then shows it through a Web interface. These data streams are provided with metadata including the position of the sensor. This can be helpful for our platform because it shows a way to structure real time data retrieval of our platform. We will need to connect real time data retrieval sensors to the platform like thermostats and weather stations and this paper will be helpful for this purpose.

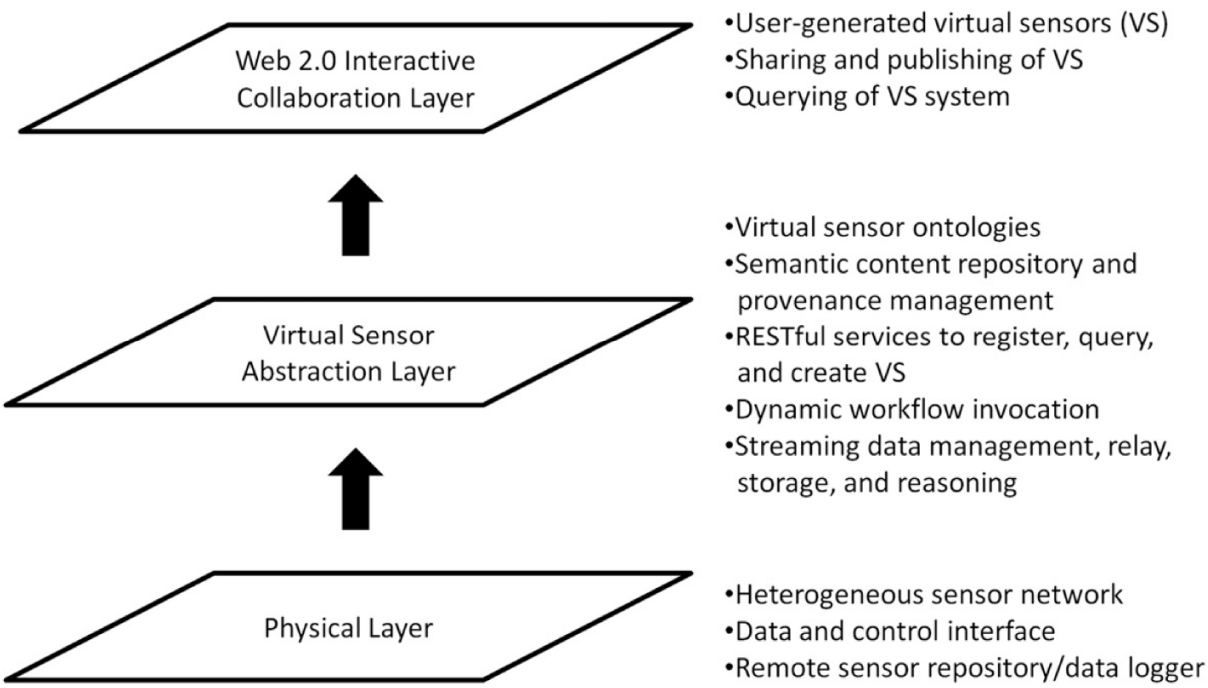

Figure 3: Architecture of the virtual sensor system

The architecture of the platform as proposed in this paper is divided in three layers. The first layer is where the remote sensor is located. It includes the sensor and the network required to connect the sensor to the internet. The authors assume that the sensors are accessible through FTP and HTTP protocols. Data is stored in the sensors and needs to be retrieved with FTP or HTTP protocols. The repositories of the public sensors are distributed in the world and they provide a list of sensors that are publicly accessible with their specification and IP address.

The middle layer is made to facilitate the retrieval, enhancement and publication of the sensor's data. This is done by implementing a "workflow-based processing (Cyberintegrator) over a semantic content repository abstraction (Tupelo), augmented by a temporal stream management layer", Hill et al. (2011). The result of this layer is a set of file-like datasets with additional information such as descriptive information and temporal relations with other datasets. It also provides a RDF statement (see section 2.2.2.2) in already defined vocabularies. These data are then accessible, for instance, through SPARQL queries.

The third layer is the Web interface where end-users are able to see the data that comes from the workflows of layer two. Visualization capabilities are provided with the Web interface such as the visualization of time-series of data that is sent through streaming by a graph.

2.2.1.3. How to Ensure the Economic Viability of an Open Data Platform

Duval & Brasse (2014) made an analysis of different business models for open data-providers. This research can be helpful for our platform because we will provide open data to external parties and customers that we will retrieve from external data sources and from the data inserted by end-users in the system. Three different business models are discussed in the paper: the free model, the freemium model and the premium model.

The first is the free model where all the income comes from advertisement (see section **Errore. L'origine riferimento non è stata trovata.**), end-users do not pay any fees to access the data and the only revenues are from advertisements that are shown to end-users.

The second business model is the freemium model where end-users are able to subscribe (and have to pay a fee) to get more functionalities of the platform. We will use this model for our platform by providing access to specific services and use cases after the end-users subscribed to a pay account (see section **Errore. L'origine riferimento non è stata trovata.**).

The third business model is the premium model where all end-users have to pay a fee to have access to the platform. This business model will be implemented for a few data sources that are requested by specific external companies. These external companies will pay a fee for every query made to our data sources using a premium business model.

The authors found that different features can be accessed by different end-users using different business models. They divided end-users in three categories: free (user is not registered, anonymous end-users), signed in (end-users that created a free account) and subscriber (end-users that created a pay account). In the free business model the free end-users have access to simple research only and they have limited access to metadata. For signed in end-users all the functionalities of the platform are provided (such as advanced search, access to API, visualization tools, social features and multilingual interface).

In the second model (the freemium model) the free users have the same rights as the free end-users in the free model but with a more limited access to data. The signed in end-users have limitations on metadata and a loss of API access. The subscriber can access the full functionalities of the system (such as API, automatic translation and full metadata).

In the premium model just the subscriber type of end-users exist, because the other types of end-users are not paying end-users and so they cannot fit in a premium business model that requires that all the end-users pay a fee to use the platform. For the subscribers all the functionalities of the system are accessible.

2.2.1.4. An open platform for personal health record apps with platform-level privacy protection

Van Gorp et al (2014) discuss a system to share health data with external applications and still maintain the privacy of that data. In the health care industry, data is difficult to share with external applications and parties because it is privacy sensitive. People do not want their privacy sensitive data to be shared with parties that in turn can share this data with other commercial companies.
The authors of this papers provided a solution to this problem. This is relevant for our platform because we have the same problem. People do not want their data to be shared and exchanged with commercial companies and they want to remain in control of their data. Using the technologies contained in this paper we will be able to provide external parties with the data of end-users and, at the same time, maintain control over this data.

Figure 4 shows the architecture of the platform as presented in the paper, called MyPHRMachines.

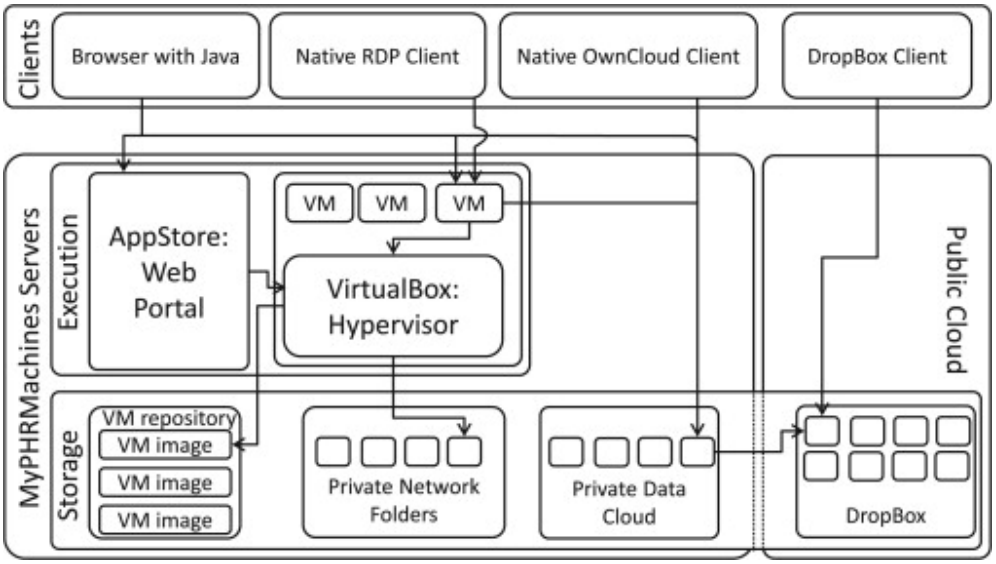

Figure 4: MyPHRMachines architecture

The architecture shows two main layers, the client layer and the server layer. The client layer includes all the methods that an end-user can use to access his data and the apps contained in the system. End-users can access the apps through a browser with Java or by a native RDP client. End-users can also access their private data through the native OwnCloud client. The OwnCloud component provides secure storage and access of the data of end-users.

In the server layer all the processing, execution and storing parts are found. The execution part, as can be seen in figure 4, is where the AppStore Web portal where the list of apps that are available is stored. The Hypervisor is a layer that includes the virtual machines that run the external applications. Every time a virtual machine is executed the end-user can access it through the Java Web interface or through a RDP client.

In the storage layer the Virtual Machines repository can be found, wherein all the images of the virtual machines are stored. Developers can upload a new version of their services and apps in this layer. Another part of the storage layer are the private network folders and the private cloud data where all the data of end-users are stored. Private network folders have better security of data because they are not connected directly to the internet. This way the virtual machines cannot exchange any information with external parties or servers because they do not have any internet connection with the outside world.

In the public cloud layer there are files of end-users that are not inside this system. End-users can connect their public cloud (defined as a cloud service with a personal account of the end-user) with the system data cloud by connecting it with the Private Data layer.

Using this software architecture the data is protected and a malicious person that wants to make a malicious app that steals information of end-users cannot do it. This is because every virtual machine is not directly connected to the internet and does not have a direct network connection. However, it is still possible to share data with other parties. If a person wants to share his data with external end-users he can provide access directly to his data through the OwnCloud client or provide access to his apps to an external person.

This platform is focused on privacy. It satisfies two main requirements: maintaining the privacy of the data and, at the same time, providing a way to use and share the data with external applications and parties.

Developers can develop their apps locally and then upload the new version of the apps in the VM repository. Every VM session is stateless. This means that every VM execution terminates when the end-user finishes to use it and then all data written on the local disk on a VM will be deleted. So developers do not need to care about the migration of running VM session. All the data that a VM wants to store must be stored in the private network folder layer, preventing any leaks of information to external parties and companies.

It is possible to use parts of this SA in our platform.

2.2.1.5. Integrating Public Datasets Using Linked Data: Challenges and Design Principles

Omitola et al. (2010) describe a method to convert datasets of different locations and formats into linked data.

The knowledge described in this article can be interesting for our platform because we will need to import data from different sources and then convert them into linked data to be able to offer them to external service providers and customers.

The authors could not take advantage of the automatic resource discovery process because the data that they wanted to import was not in the linked data format. So they sourced their datasets by looking source by source and site by site. The data that was found was in many different formats such as PDF, HTML and XLS.

It was decided to use RDF (see section 2.2.2.2) as the normal dataset form and the turtle representation of RFD triples. They chose the 4store system to manage and store RDF triples. Furthermore, the Exibit publishing framework (see section 2.2.2.5) was used to develop the interface for the consumption of data.

The paper includes the complete step-by step process that was followed to convert the different types of data to linked-data. First of all they modelled the datasets and found links with existing well-known vocabularies to re-use terms. Then the datasets were converted to the RDF format. The problem of this step was the little semantic description of the data involved. In some of the public data there was no explanation of the relationship with other data. Another problem was that it was difficult to integrate and link data from disparate data sources.

They developed scripts that automatically converted the spreadsheet data into RDF triples. They also used the Jena Semantic Web Framework (see section 2.2.2.6) to convert types of data. Then they aligned the different datasets by aligning the different identification of relation between the different relevant entities of the datasets. Finally they developed the interface for the linked data consumption.

2.2.2. Existing technology solutions

In this section we will analyze and describe solutions using existing technology that are related to our platform. We found these technologies by searching on the internet for the main technologies related to our platform. The main search engine that we used was Google.

In Figure 5 below all the different existing solutions that are relevant to our platform are shown.

The Linked data section shows different technologies such as RDF, SPARQL and Linked data fragment technologies.

In the Linked data framework section there are the two frameworks discussed in this section: the Exhibit framework and the Apache Jena framework.

In the Identity access framework section two different identity frameworks are described that we may implement in our platform. Both the Linked data framework and the Identity access framework technology may be implemented in the software architecture of our platform.

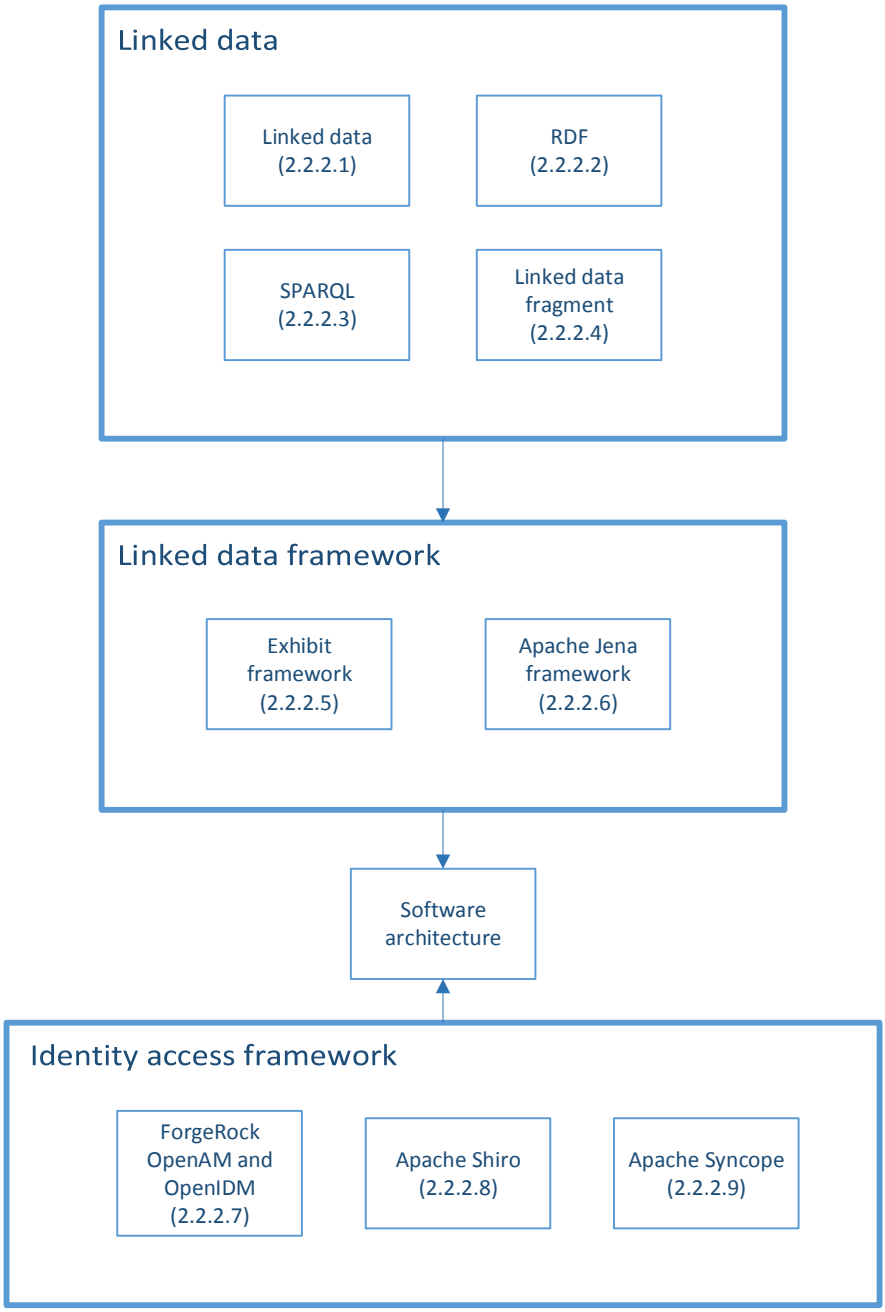

Figure 5: Existing technology solution schema

2.2.2.1. Linked data

Linked data refers to using the Web to make typed links between data from different sources. Technically speaking, Linked data refers to data that is readable from a machine and that is published on the Web (Bizer et al., 2009)

On the Web, hypertexts are programmed using HTML documents. HTML allows the programmer to connect different documents together by using hyperlinks. In Linked Data there is RDF (Resource Description Framework). The RDF defines how different objects link together. In the next section a more accurate definition of RDF is provided. Also in Linked data every object or concept is identified by an URI , this is similar to the Web (Berners-Lee, 2006). An URI (uniform reference identifier) is a string of characters that refers to a resource.

Linked data, following Tim Berners-Lee, should follow rules (Berners-Lee, 2006). The following rules are the original rules of Tim Berners-Lee:

1. Use URIs as names for things.
2. Use HTTP URIs so that end-users can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF*, SPARQL).
4. Include links to other URIs. so that they can discover more things.

Figure 6: Tim Berners-Lee's rules for Linked data

Open data can be viewed also as a browsable graph. With a browsable graph a person can see in a graphical way the connections of every node. When one looks for a node in the RDF graph, the server returns all the information about the arcs in and out of that node. Thus it is possible to know all the statements where the node is a subject or an object and the description of all the blank nodes attached to the node. (Berners-Lee, 2006)

Tim Berners-Lee also made an analysis of the different ratings of Linked data. These levels can be used for Linked Open Data and for Linked Data. The only difference is that Linked Open Data provides data that is accessible under an open license.

Figure 7 shows the ratings, starting from the lowest (1) to the highest (5) (Berners-Lee, 2006).

1. Data is available on the web (in whatever format).
2. Data is available as machine-readable structured data.
3. Data is available as machine-readable structured data plus non-proprietary format.
4. Data is available as machine-readable structured data, non-proprietary format plus uses open standards from W3C (RDF and SPARQL) to identify things.
5. Data is available as machine-readable structured data, non-proprietary format, uses open standards from W3C (RDF and SPARQL) to identify things plus data is linked to other users' data to provide context.

Figure 7: Berners-Lee's ratings of linked data

2.2.2.2. RDF

The RDF (Resource description framework) is a data model for publishing structured data on the Web. (Heath & Bizer, 2011). The RDF data model can represent information with a node and an arc labeled directed graphs. The objective of RDF is to moderate between other data models used on the Web (Heath & Bizer, 2011).

With RDF, the description of a node is represented with triples. The different parts of each triple are called subject, predicate and object. A triple can be seen as a basic sentence structured, for instance:

"Diego is a student"

Becomes:

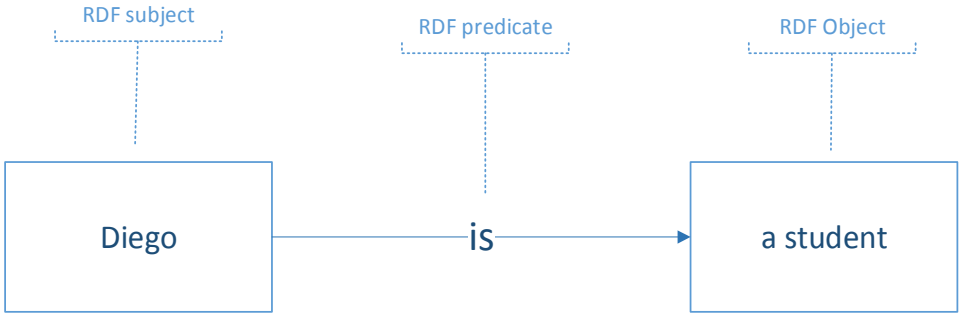

Figure 8: RDF triple structure

Each part of a triple can be a URI that connects to an external object. Each part of a triple can also be a literal value, like a number, date or string. The predicate indicates which relationship exists between the subject and the object. A predicate can be also a URI and can be taken from vocabularies, those are collections of URIs that contain predicates about a specific domain.

The following example converts a sentence to a RDF graph.

Suppose that we have the sentence:

"Diego is a student and works as a business consultant"

The **subject** is "Diego".
The **objects** are "student" and "business consultant".
The **predicates** are "is" and "works as".

Then the according graph will be:

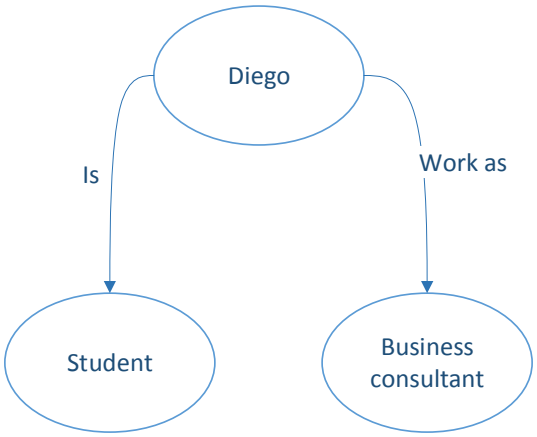

Figure 9: RDF graph for a sentence with to objects

2.2.2.3. SPARQL

The SPARQL language is used for querying RDF data. It enables the user to retrieve and manipulate data stored in the RDF format. The current last version of the language is the 1.1 version that was released in March 2013. ("SPARQL 1.1 Overview,")

In the following example taken from the W3C SPARQL Working Group ("SPARQL 1.1 Query Language,") we will make a SPARQL query to retrieve the title of a book.

Suppose to have this data:

```
<http://example.org/book/book1> <http://purl.org/dc/elements/1.1/title> "SPARQL Tutorial" .
```

Figure 10: RDF data

And to have this SPARQL query:

```
SELECT ?title

WHERE

{

 <http://example.org/book/book1> <http://purl.org/dc/elements/1.1/title> ?title .

}
```

Figure 11: SPARQL query

The select clause identifies the data that will appear in the results, and the where clause provides the RDF data that needs to match.

The result of this query will be:

Title
"SPARQL Tutorial"

Figure 12: SPARQL query result

2.2.2.4. Linked data fragments

Linked data fragments (LDF) (http://linkeddatafragments.org/) is a method that needs minimal server effort during the solving of SPARQL queries and that allows to switch the effort of the queries to the client. This method was introduced because public SPARQL endpoints that reply to SPARQL queries are often unavailable due to the high numbers of queries that they need to reply to.

One solution to this problem was to create local endpoints using data dumps, but the data of the local endpoint was not up to date and the client could not query every dataset but just the local data of the local endpoint. To solve this problem the LDF was invented, making it possible to switch the effort from the server to the client.

The Linked Data Fragment of a Linked Data dataset consists of the triples of the dataset that match a specific selector. A LDF server consists of a HTTP server that offers the service of LDT with one or more of its datasets. The LDF technology can be installed in any HTTP server that wants to offer this service.

The next diagrams compares the data fragments solution with the normal SPARQL data retrieving and with the local endpoint data dumps.

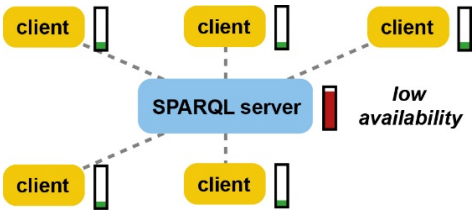

Figure 13: SPARQL normal querying

Figure 14: Local endpoint data dumps

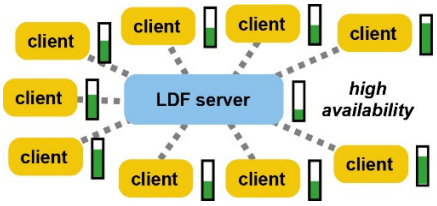

Figure 15: LDF technology

As is shown in the above figures, the normal SPARQL querying relies on the server for the querying effort. The local endpoint data dumps relies on the local server but the data may not be updated. With the LDF technology the server needs to put less effort in the query and the clients do a big part of the effort.

2.2.2.5. The Exhibit publishing framework

Exhibit (http://www.simile-widgets.org/exhibit/) is a publishing framework to integrate data in web pages. It enables incorporating tabular data into webpages using simple tags to incorporate in the HTML code. Then the framework automatically imports the linked data from the tabular data sources. To attach the data a developer just needs to specify the data columns that contain coordinates, dates or attributes and then they will be visualized in the web interface.

The framework does not need installation because it is an open source JavaScript library that relies on the client browser. So it can be integrated in already existing CMS in a simple way.

This framework may be used for our platform.

2.2.2.6. Apache Jena

Apache Jena (https://jena.apache.org/) is an open source framework to build linked data and semantic web applications. The framework is composed by different APIs that process RDF data.

This framework permits users to develop linked data applications in a faster a simpler way by offering APIs that can be used in the development of the software architecture of a platform.

It offers TDB, which is a component for RDF storage and query. It can be accessed and managed with scripts and by using the Jena API. It can be used between multiple applications by the Fuseky component that a SPARQL server provides.

The ARQ is the query engine of Jena. It supports the SPARQL RDF query language and many other functions such as free text search, access of the SPARQL algebra, custom filter functions, property functions for custom processing of semantic relationships, aggregation, GROUP BY and assignment as SPARQL extensions. ("ARQ - A SPARQL Processor for Jena,")

It implements an advanced text function that allows making free text searches within SPARQL queries. ("Text searches with SPARQL,")

We will be able to use this software in the SA of our platform.

2.2.2.7. ForgeRock OpenAM and OpenIDM

OpenAm (https://www.forgerock.com/) is an open source identity provider that manages the identity of end-users. It offers a modular platform where end-users can store their login details, their policies and their details. It provides easy integration with external services and web platforms.

With this platform end-users are able to securely manage the access to their services. It provides a personalized experience on any digital channel, such as mobile, cars, home appliances or anything connected with the internet. The authentication is based on context access that includes location, IP address and time of the day. Its policy engine tool permits managing access to the different services that an end-user can have access to. It eliminates the need to use multiple passwords to use different services by providing a unique identity manager. It allows a rapid customer adoption by permitting quick integration options such as the "sign up and log in with Facebook" function.

OpenIDM provides a complete identity management platform that supports following the customer's login in every device and service. It permits an end-user provisioning based on existing rules and permissions. End-users can manage registration and password reset themselves without the help of an external system administrator. It also allows to connect the identity of end-users with cloud services like Google and Office365.

2.2.2.8. Apache Shiro

Apache Shiro (http://shiro.apache.org/index.html) is a Java security framework that has the following functions: authentication of end-users, authorization, cryptography and session management. Apache Shiro provides easy to use APIs to provide a secure authentication to a web application.

Using Shiro, managing security rules for every end-users registered in the system becomes simple . It permits to react to events during the different phases of end-user's authentication. It also allows to aggregate more data sources of end-user security data and to present this data with a single composite view. It also offers the "remember me" function for end-user access without login.

Data stored with Shiro can be cryptographed to increase its security. It supports the "run as" function that allows an administrator to login with the credentials of an end-user This can be useful in administrative scenarios.

We may be able to use this software in the SA of our platform.

2.2.2.9. Apache Syncope

Apache Syncope (http://syncope.apache.org/) is an open source system that can be used to manage digital identities in enterprise environments. It provides different functions such as account and role administration, domain integration with external resources and a web-based interface for the administration console.

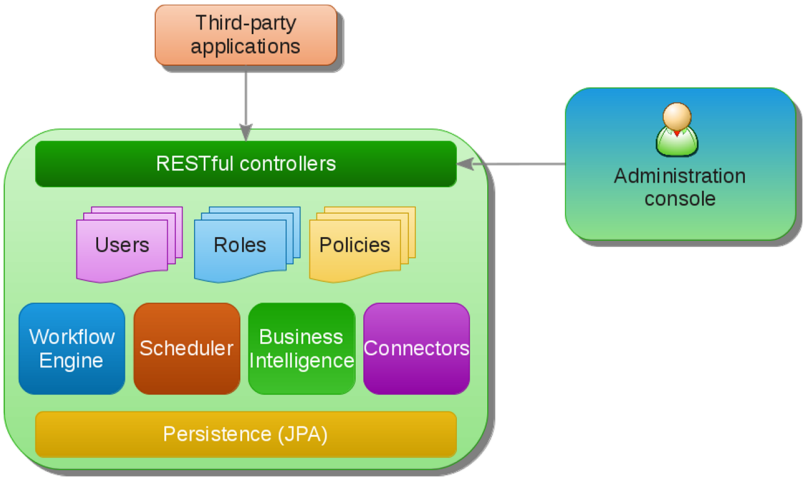

Figure 16: Software architecture of Apache Syncope

The figure 16 shows the SA of Apache Syncope. The core part is formed by the web application that implements the IdM features. The central component manages the data flow through the system. It handles RESTful call, processes data and propagates it to and from external data sources.

The administration console allows managing the identity of end-users on the web interface.

The connection layer connects the IdM part with the other parts of the SA of our platform.

We may be able to use this software in the SA of our platform.

2.2.3. Products and services on the market

In this section organizations that are already developing or researching a data-collecting platform are analyzed and discussed. In addition to our own search results several experts gave us useful information about organizations that are of interest to our project.

For each company we state its name, web address, current status of development, its description and its current provided functions. This analysis was made in January and February 2015 so the information provided in this section is fairly up to date.

2.2.3.1. My data store

Company

Telecom Italia Group

Website

http://jol.telecomitalia.com/jolskil/personal-data-store/

Status:

Prototype testing with MTL community members

Description:

Telecom Italia is creating an aggregator of Linked Data that will permit end-users to decide their privacy setting for each data source.

Functions:

- Collecting of personal information: end-users can collect data regarding themselves. End-users can decide to collect particular data just during specific time-slots.
- Sharing of information: end-users are able to share their personal information with third parties. End-users are able to set different disclosure levels for each piece of information.
- Deleting of information: end-users are able to delete information that they do not want to have anymore.
- Interface for aggregating information: information is aggregated into a single interface that permits end-users to have a complete view of their information. This is done to increase the end-users' consciousness of daily behaviors.
- Compare data with other end-users: end-users are able to compare the data that they share with data shared by other end-users.

2.2.3.2. The HAT project

Company

HAT Research Team

Website

http://hubofallthings.com/

Status:

First beta launch in June 2015 in UK and Singapore. They are collecting marketing data through the mandatory survey to register to the service on the website.

Description:

The HAT project team is developing a platform to collect Linked Data. Data can be shared with external applications. People will be able to set privacy settings for their data and to analyze their data. External companies will be able to set-up their HAT servers and to store personal data of end-users inside them. The HAT project will

validate external apps, maintain a catalog of the external HAT servers, be an intermediate gateway for inter-platform payments, manage the APIs, manage the universal HAT unique ID database, install and update the external HAT servers and give advice about business and revenue models to external companies.

Functions:

- Collecting of personal data: end-users can collect and store data that comes from different sensors, devices, internet connected objects and services. End-users can store this data in a HAT server.
- External applications and services: data can be shared with external applications and services that will use and analyze it.
- End-users will be able to see in one integrated interface the apps that they can use with a specific HAT server.

2.2.3.3. CommonSense

Company

Sense Observation Systems

Website

http://www.sense-os.nl/home

Status:

The beta version of the platform and of the mobile app have already been developed. There is the possibility to have free and open access to a beta account to try the platform.

Description:

CommonSense provides a platform where data from different input sources is stored. Data can be retrieved from the mobile app developed by the company and from a Fitbit device. The system can retrieve and store the following types of data:

- Sleep information
- Time activity information
- Location information

End-users are able to share their data with other end-users.

Functions:

- Collecting information: end-users can collect information using the mobile app, Twitter and a Fitbit device.
- Presentation of information: end-users can see their information on the dashboard.
- Sharing of information: end-users can share their information with other end-users and see information shared by other end-users.
- Open API: external companies and developers can build apps for the platform using the CommonSense's APIs.

2.2.3.4. Tippiq

Company

Tippiq (Alliander)

Website

https://www.tippiq.nl

Status:

The first release is public and everyone can access and register to it.

Description:

Tippiq provides a platform to find information about events, projects and initiatives of an area.

End-users can search with their postal code to information about their area and register to the system.

Functions:

- Presentation of information: end-users can see information about their surrounding on the dashboard.
- Inclusion of information: end-users can provide information about their area to the system that can be shared with other end-users of that area.
- Privacy of information: end-users are able to decide with whom to share their information.
- End-users are able to set up an initiative in their neighborhood and share it with other Tippiq end-users.

2.2.3.5. Huislink

Company

Huislink

Website

https://www.huislink.nl/

Status:

The first beta release is public and everyone can request an account.

Description:

Huislink provides a platform to store all information regarding a house. For instance photos, characteristics and general information about a house. This data can be shared with other end-users that can search the houses database to find a house to buy or rent.

Functions:

- Storing of information: end-users can store information regarding their house on the platform.
- Sharing of information: end-users can share their information with other end-users.
- Searching of information: end-users can search for information regarding houses in the system and they are able to contact the owner of the house to rent or buy the house.
- Messaging system: end-users are able to add other end-users to their profile and to exchange messages with them.

2.2.4. Current problems and limitations of existing products and services

Currently a solution that provides a platform that satisfies all our stakeholders' needs and requirements (as discussed in section 2.1.2) does not exist. There is a gap between what we need and what has already been developed.

In the next table a comparison is made between different solutions available with the requirements that are described in section 2.1.3. The requirements list formed the bases of this comparison and we compared it with the functions offered by each existing solution. For a better description of each requirement please refer to section 2.1.3. A blank spot in the table means that a requirement is not implemented for the specific solution.

Requirements/existing solutions	My data store	The HAT project	CommonSense	Tippiq	Huislink
Different levels of authentication for end-users					
The possibility to import external data	Present	Present	Present	Present	Present
The possibility to store data inserted by end-users					Present
External services development		Present	Present		
Dashboard	Present just for the aggregation of information	Present without the newsfeed	Present without the newsfeed	Present without the newsfeed	Present without the newsfeed
News feed					
End-users receive alerts					
Social function					Present
End-user function to share their data with external websites with a wallet function					
Tablet and mobile browsing			Present		

Table 1: Requirements and technologies comparison

Some of the requirements are met but in none of the platforms analyzed all the requirements are met by one single platform. Moreover, several of the mentioned platforms do not focus on data specifically regarding houses but they focus on different types of data.

3. The business case

3.1. External view

In the external view (see Figure 17: External view) there are three different main entities that will link to our platform. The first group consists of the service providers that will provide the services in the platform (the complete description is in section **Errore. L'origine riferimento non è stata trovata.**). The second group is made up by the data-providers, they will provide the data sources that will connect to the platform (the complete description is in section **Errore. L'origine riferimento non è stata trovata.**). The third group are the end-users. They will use the system, its services, functions and data sources (the complete description is in section 0).

A service provider can also be a data provider (and vice versa) because every company or organization can both share data with our platform and offer services that link to these data. An end-user can also be a data provider by entering his personal data in the platform.

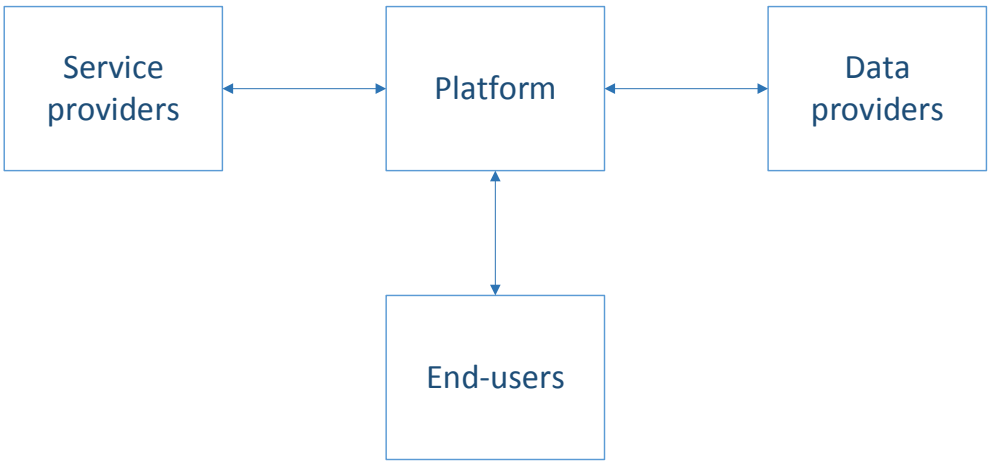

Figure 17: External view

3.2. End-user perspective services

This section has been left out because it is part of the private version of this thesis.

3.3. Use cases summary

This section has been left out because it is part of the private version of this thesis.

3.4. Most important use cases

This section has been left out because it is part of the private version of this thesis.

3.5. Data sources

In this section the data sources that will be used in the use cases that are implemented in the first release of the platform are presented (as described in section 3.1).

These data sources were found by analyzing different use cases, as described in section 3.1, and by finding out which data sources were needed for each use case.

In Appendix A a list of the data sources that will not be implemented in the first release of this platform is shown, and that are not connected with any use cases described in section 3.1. A person can use these data sources for future work and for future development of new use cases.

In the next table the meaning of different sections of the Table 3: Data sources is explained.

Name	**Description**
ID	Sequential ID number of the data source.
Name	Name of the data source.
Description	Description of the data source.
Category	Category of the data source. This category will be used to group the different data sources in the end-user interface.
How to retrieve it	How to retrieve the data source.
Use case that uses it	The corresponding use case that uses the data source.

Table 2: Data source list explanation

In the next table the different data sources are listed.

ID	**Name**	**Description**	**Category**	**How to retrieve it**	**Use case that uses it**
1	Alarm data	Information about the alarm (the current state of the alarm)	Security	From the alarm	Security alarm sharing

2	Old photos of the surroundings	Old photos of the surroundings of the house where the end-user lives	Neighborhood	Archief Eemland photo database	Photo retrieval and upload
3	Companies in the surrounding area information	Information about companies that are in the surrounding area of the end-user's home	Neighborhood	Dutch chamber of commerce	Chamber of commerce company information
4	Information about appliances	Static and dynamic information about appliances	Various	Appliance producers' datasets	Maintenance and various information about appliances
5	Security information	List of criminal offences in the area	Security	Burgernet datasets	Security information
6	Energy information	Details about the end-user's energy consumption regarding gas, water and electricity	Energy	Energy providers' datasets	Energy information
7	Flood risk information	Risk of flood in the end-user's area	Security	Government datasets	Flood risk information
8	Garbage collection information	Garbage collection timeline of the end-user's area	Municipality	Government datasets	Garbage collection information
9	Municipality contact information	Contact information about the end-user's municipality	Municipality	Municipality datasets	Municipality contact information

Table 3: Data sources used for the first release of the platform

3.6. Revenue models

This section has been left out because it is part of the private version of this thesis.

3.7. Foundation model and open source model

In this chapter we the advantages and the disadvantages of creating a foundation and of developing the platform as OSS are described. The HuisKluis foundation already uses this model (for further information about HuisKluis please refer to the preface of this thesis).

3.7.1. Foundation model

3.7.1.1. Description

Using a foundation instead of a private company we will be able to gain more funds to develop the platform. Since a foundation is not allowed to make any profit and share it with the shareholders, the profit for the creators of the platform will come from other sources, as is explained in section 3.7.1.4.

A foundation normally consists of a foundation board and employees that work for the foundation. The foundation board normally consists of end-users that do not work as employees for the foundation, so we will not be members of the foundation board. The members of the foundation board will be end-users from external parties that fund the foundation or other external parties.

3.7.1.2. Dutch foundation principles and regulations

In The Netherlands foundations have a great history. In 2001 the non-profit sector in Netherlands had about 669.000 full-time equivalent jobs that added up to 12.9% of non-agricultural paid employment. The amount of volunteers in the non-profit sector in Netherlands in the same year was about 400.000 full time equivalents jobs (Burger & Dekker, 2001).

Foundations are composed of two types of end-users, the board of directors and the employees. They have different key activities in the foundation as is explained in section 3.7.1.3. A foundation cannot share profit with its members but must use them for its purpose and mission. When a foundation is registered, several things need to be established: the objective of the foundation, the method of acquiring new board members and the resignation procedures of existing ones (Burger & Dekker, 2001). The foundation has full legal competence and the management board's end-users are personally responsible for legal acts of the foundation (as long as the foundation is not entered at a commercial register) (Burger & Dekker, 2001).

The Dutch chamber of commerce requires an annual fee for the registration of the foundation (Burger & Dekker, 2001).

The management board is responsible for realization of the objectives and for representing the foundation. It is able to take independent decisions. There is the possibility to have an supervisory board (especially in large foundations) that supervise the actions of the management board (Burger & Dekker, 2001) but this is not mandatory. As a rule, the management board's directors cannot have any conflict of interests with the foundation.

In Netherlands just a few types of foundation can have privileged tax treatment and this is correlated with the type of services and activities that they offer. Foundations that pursue "religious, ideological, charitable, cultural, scientific or public interest objects" are eligible for this tax reduction and they are called "public benefit organizations". (Burger & Dekker, 2001)

A common way to use the foundation model in Netherlands nowadays is to create a foundation that owns the shares of a company. The foundation board has the right to vote for the decisions of the company whose shares have (Francken, 2013).

3.7.1.3. Foundation board and employee key activities

The foundation board will have the following responsibilities, as is explained in the next table.

Title	Description
Coordinate consulting services	The foundation board will coordinate consulting services that the employees do for external companies. There will be the possibility that external end-users and employees offer private consulting services to external companies regarding this platform. In this case, the foundation board will coordinate just the relationship with them and will not coordinate their work.
Coordinate funds retrieval	The foundation board will coordinate funds retrieval from external parties, both public and private.
Coordinate partnerships	The foundation board will coordinate partnerships with external companies that provide services and data sources to the platform.
Coordinate the ethical compliance of all parties involved	The foundation board will constantly check that all different parts of the foundation, including the employees and external partners, will comply with the ethical principles of the foundation's manifesto.
Coordinate marketing activities	The foundation board will coordinate the marketing and public relationship activities.
Coordinate employees	The foundation board will coordinate the main activities of employees of the foundation.

Table 4: Foundation board activities

The employees of the company will have the following tasks and responsibilities, as listed in the next table.

Title	Description
Managing the open source community	The project managers of the foundation will manage the open source community that develops the software.
Defining new business models for the platform	The economists of the foundation will search for new businesses and for new revenue models for the platform that will benefit all parts involved.
Providing business and strategic advice	Manager will provide business and strategic advice to the foundation board regarding the platform
Managing public relationships with external parties	The public relationship department of the foundation will manage the public relationship with external parties.

Managing the marketing of the foundation	The marketing department of the foundation will manage marketing of the platform.
Analyzing future functions, services and data sources of the platform	The innovation department will manage new functions, services and data sources that the platform can implement in the future.

Table 5: Foundation employee activities

3.7.1.4. Revenue

In the foundation model we will work for the foundation as employees or as foundation board members and for external companies as private consultants.

As employees of the foundation we will get a normal hourly rate. As private consultants we will be able to provide the services explained in the following table.

Title	**Description**
Consulting services to companies that want to connect their data sources to the system	We will provide consultation for external companies that need to connect external sources to the platform.
Consulting services to companies that want to provide services linked to the system	We will provide consultation for external companies that need to make services linked to the platform.
Consulting services for companies that want to offer a personalized version of the software to their customers.	We will provide consultation for external companies that need to make a personalized version of the platform to offer to their customers.
Consulting services for the foundation	We will provide consultation to the foundation as external consultants for various activities such as managing and innovation.

Table 6: Consultancy services

In addition to the revenue models described in section 3.6, the foundation will also receive funds from the revenue models explained in the following table.

Title	**Description**
Private and public donations	People will be able to make donations to the project to support it. The donations will be managed by the foundation.
Certification of platforms implemented by other companies	If an external company wants to implement a copy of the platform and to have certification that its platform is compliant with the original project, it will need to pay a fee to the foundation.

Table 7: Foundation additional revenue models

3.7.1.5. Foundation manifesto

The foundation will follow ethical rules that are explained in the following manifesto.

The foundation board will always check that all parts of the foundation comply with these rules.

1. In the last 30 years the world changed into a digital world. We will help end-users to take full advantage of this digital world by providing them with useful and simple services.
2. Strangers cannot enter your house uninvited, it is prohibited by law. But big technology companies can record your data through your telephone, television and computer and use it for economic purposes. We will solve this problem by permitting you to share your data with just the external parties that you want.
3. Every Dutch person has the right to privacy. They have the right to choose with whom to share their data. Currently this is not possible, big technology companies use data of end-users, who do not have any control over them. We will give end-users the right of privacy in the digital world back.
4. National governments collect private data of citizens for anti-terrorism purposes. We will share data of end-users with national governments just if the national law forces us to share data with them. We will make a complete and transparent list of situations where we are forced to share data of end-users within the national government and we will share end-users' data just in these specific cases.
5. People must have the right to change their private data when they want. We will help end-users to change their data across the internet with any measure that we will be able to do.
6. We will only be successful when we reach millions of end-users. To do this we will provide easy, safe and useful services to end-users.
7. Open source software changed the world in many ways. Open source software is transparent, safe, reliable and innovative. The platform that we will make will be open source so end-users are able to benefit from the advantages of the open source system.
8. Some services on the internet are privacy sensitive, such as online banking. We will provide high security measures that control the identity of end-users that access highly private services.
9. To be really innovative, ideas and opinions of all parts involved in the project will be taken in consideration. We will carefully listen to every idea, concern and opinion that is voiced by our customers and partners to deliver the best innovative solution that will satisfy the needs of all the end-users involved.
10. Real innovation is created by honest and trustworthy actions. We will put all our statements in action by providing a platform that is safe, private, secure and simple, to satisfy the needs of end-users from all over the world.

3.7.2. Open source model

It will be possible to develop this platform as open source.

The objective of an open source strategy is to create an environment where open-source developers will develop the project without receiving a salary for developing the code. Following this strategy, the complete project will be released under a GPL license permitting anybody to obtain and modify the source code. Any modification of the source code, following the GPL license ("GNU GENERAL PUBLIC LICENSE," 2007), must be made public. There are also other types of licenses for OSS that have more or less restrictions than the GPL license. In this section we assume that we will use the GPL license.

Using this strategy, we will create a foundation that will own the platform brand, logo and domain. The foundation will coordinate the activities of the open source developers and will provide business, marketing and management activities.

3.7.2.1. Advantages and disadvantages

The advantages of implementing an open source model are explained in the following table (Morgan & Finnegan, 2014).

Title	Description
Free developers	Most of the developers will work for free. This will lead to a big cost saving for us. We will still need to pay the management and leaders of the project that will coordinate the open source developers. There is the possibility that we will still need to hire developers that will work on the project if the reaction of the open source community to our platform will not be enthusiastic.
More trust from customers	Open source software is perceived to be safer, more reliable and trustworthy by customers, so they will have more trust in our platform. Customers will not have the vendor lock-in because in a few years there will be copies of our platform in the market and customers will be able to switch to our competitors without any problems.
Better quality of code	Due to a better peer review and because everybody will be able to access the code and report bugs, the quality of the code will increase. This will lead to having a more reliable and trustable product.
Simple for third parties to develop services and plugins for the platform	Since third parties will be able to access the source code, it will be simpler for them to develop services that connect to the platform.
Accelerate innovation and exchanging ideas	A lot of developers will be able to see the source code. Because of this, they will be able to implement and propose innovations and ideas to the project.
Access to superior knowledge	Since a lot of developers will be involved in the project, they may have better skills and knowledge of developing compared to internal paid developers.

Table 8: OSS advantages

The disadvantages of implementing an open source model are explained in the following table (Morgan & Finnegan, 2014).

Title	Description
Difficult to find OSS developers	According to Santos et al. (2013), for a new project like this it can be difficult to attract new open source developers in the initial stages of development due to the low attractiveness of this project. The low attractiveness to the OSS environment stems from the fact that this is a new project.
Difficult to manage the creation of an open source software	Because developers will work for free, we will not be able to control their work and schedule. We will also need to take their opinions into consideration during the decision making process. This will lead to a lack of centralized decision making. The overall cost of managing this may overshadow the initial cost saving of using free developers.

No ownership of the code	The code can be used by anybody. The source code is one of our main assets and, with open source developing, the entry barriers to future competitors will be low because they will be able to use our source code. This may lead to less revenue and profit.
Difficult to access top quality developers	Because developers will work for free, the top quality developers may not join our community because they are already developing for closed source companies and getting paid for it.
Difficult to plan a roadmap	Due to the difficulty of managing the community it will be difficult to schedule a roadmap and to force developers to follow it.

Table 9: OSS disadvantages

4. Software architecture

In this chapter the software architecture of the platform is described. The reader can find the requirement analysis, the technical problem analysis, the solution domain knowledge analysis, the solution analysis and then the synthesis of the overall software architecture in this chapter.

4.1. Software architecture design method

The software architecture design method that we will follow is described in Figure 18. The design is based on the process explained by Tekinerdoğan and Akşit in their book "Synthesis-based software architecture design" (Tekinerdoğan & Akşit, 2002). This design process involves the construction of sub-solutions for every different problem of the architecture and then the integration of these sub-solutions into one single solution.

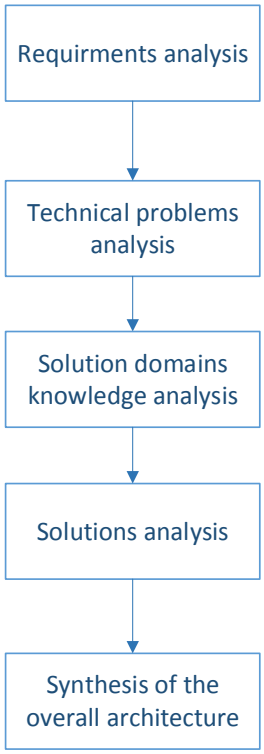

Figure 18: Software architecture design method

As explained in Figure 18 the first phase will be the "requirement analysis" phase. The requirement analysis consists of a list of the requirements that the platform needs to have from the client's perspective.

Then, in the "technical problem analysis" phase, the client's requirements are translated into technical requirements. Normally the client's requirements lack technical specifications and the right details of the problem. In consequence, the gap between the client's requirements and the software architecture is generally large.

In the "solution domains knowledge analysis" phase, the solution domains for each problem are mapped. Each solution domain is connected to a technical problem in the previous step. This phase is different from the "solution analysis" phase because "the solution domain itself may be large and include abstractions that are not relevant for solving the corresponding problem" (Tekinerdoğan & Akşit, 2002). This phase enables creators to have a better distinction between and better knowledge of the solution domains.

Then, in the "solution analysis" phase, the solution analysis for each technical problem is provided. The solution analysis includes the part of the software architecture that will solve that specific technical problem.

In the "solution analysis" phase the commonality and the variability analysis of the alternatives that can be developed for a solution domain are also included. During this phase alternatives of the same solution domain are analyzed and then the commonalities and the variability between them are found. Finally several of these variables and commonalities are taken into consideration to decide on the final solution for that particular technical problem. In this thesis the commonality and the variability analysis is done just for technical problem number 11: "what will be a suitable architecture for our system that provides a secure and reliable system and that will arrange all the components within the system?". This is because it is the main technical problem of the software architecture.

In the "synthesis of the overall architecture" phase the different solutions of the previous phase are synthesized in the overall architecture of the platform.

4.2. Requirement analysis

This section has been left out because it is part of the private version of this thesis.

4.3. Technical problems analysis

This section has been left out because it is part of the private version of this thesis.

4.4. Solution domains knowledge analysis

This section has been left out because it is part of the private version of this thesis.

4.5. Solution analysis

This section has been left out because it is part of the private version of this thesis.

4.6. Synthesis of the overall architecture

This section has been left out because it is part of the private version of this thesis.

4.7. Scenario analysis

This section has been left out because it is part of the private version of this thesis.

5. Mockups of the UI

This chapter has been left out because it is part of the private version of this thesis.

6. Software prototype

This chapter has been left out because it is part of the private version of this thesis.

7. The trust issue

This chapter has been left out because it is part of the private version of this thesis.

8. Discussion

This thesis contains a design for a platform to provide home based services concerning care, safety, comfort and fun. This section will analyze the results that were achieved during this project.

The first result that we obtained is a list of requirements of the different stakeholders that are involved. These requirements were taken from interviews with the stakeholders. The interviews that we conducted demonstrated that several needs of the stakeholders are currently not met and that our platform can fulfill them. The research that we did to find these requirements was extensive, but could be done more thoroughly. The number of end-users that have been interviewed is low and further research on this should be done to confirm these requirements. Further research can also be done using the prototype described in chapter 7 to test if the requirements are really needed by the different stakeholders involved.

Regarding the existing available technological solutions and the current products and services on the market, most of them were analyzed in chapter 2 and were integrated in the technological solutions of the SA of the platform. However, more of them could be analyzed to have a wider spectrum of solutions to implement in the platform and to know more about our current competitors.

In section 3.3 a summary of the possible use cases to implement in the platform was made. This list contains 55 use cases and they are briefly described in that section. However, more use cases can be found by deeply analyzing the needs of the stakeholders and by showing them the prototype described in chapter 7. This will lead to an even more extensive list of possible use cases.

In section 3.4 nine use cases are deeply analyzed. We made a deep analysis with sequence diagrams for each of them. However, they are a low number. More use cases should be thoroughly analyzed in the future.

The data sources described in section 3.5 are strictly connected to the use cases described in section 3.4. In Appendix B other data sources are mentioned that can lead to more future use cases. In the future, new data sources should be analyzed to be implemented in the platform.

The software architecture described in chapter 4 contains and satisfies all the requirements mentioned before. However, we did not test this SA by implementing it. When it will be implemented, maybe future research will need to be done in order to improve it.

In chapter 5 the mockups of the UI are shown. This UI was tested by two professional user interface graphic designers. However, the real test will be when the platform will be developed and when it will be tested by the end-users. Because of this, more research on the UI of the platform will need to be done. The same advice should be given about the software prototype described in chapter 6. It was tested with 5 possible end-users to check if the requirements that it is supposed to satisfy have really been satisfied. The results were positive overall but the real test will be when the platform will be created and more end-users will be able to use and test it.

The solution to the trust issue described in chapter 7 is innovative but it still needs to be tested in the real world. When the platform will be created we will be able to test if our trust solution will lead to end-users sharing their private data with us and the services of the platform.

9. Summary of the conclusion for the public version of the thesis

This chapter contains a summary of the conclusion of this thesis. The complete conclusion of this thesis is available on the private version of this thesis.

In this chapter conclusions about the project are drawn. The different research questions stated in chapter 1 are answered. The scientific contributions, both theoretical and practical, are stated and described. Finally, the research limitations and the recommendations are provided.

In this chapter conclusions about the project are drawn. The different research questions stated in chapter 1 are answered. The scientific contributions, both theoretical and practical, are stated and described. Finally, the research limitations and the recommendations are provided.

9.1. Reviewing the research questions

In this section we will review the research questions that we posed in chapter 1.

- **RQ1: What are the characteristics a platform where end-users will be able to see, share and analyze information regarding their houses?**

This research question is quite broad and has few limitations on scope. For this thesis a platform was developed where end-users will be able to see, share and analyze information regarding their houses. We made a problem analysis to find out which requirements the platform will need to have by analyzing the needs of the stakeholders involved. This analysis was made by interviewing different parties involved and by collecting their needs, opinions and requirements.

In chapter 3 the reader can find the complete business case of the platform. The end-user perspective services are provided with a complete description of each service.

- **RQ2: Which use cases will be implemented that use information about houses and that will provide useful functions to end-users?**

In section 3.3 a complete overview of the different use cases that can be implemented is provided. This list contains 53 use cases and provides a brief description and categorization of each of them.

In section 3.4 several use cases are deeply analyzed. These use cases were the most useful for the different parts involved and they will be implemented in the first release of the platform.

For more information please refer to chapter 3 of this thesis.

- **RQ3: What will the software architecture for such a system look like?**

The SA of the platform is completely explained in chapter 4. The SA that we developed is described in detail in section 4.6 and it is divided into several parts.

- **RQ4: What will be the solution to the trust issue between end-users and the information sharing platform?**

End-users will need to trust our platform to store their private data.

For further information please refer to chapter 7 of this thesis.

9.2. Research limitation

A number of limitations of this research are present.

The most important limitation of this project is the small sample size of the interviews of the end-users described in Appendix A. Due to time constraints and practical limitations we could only work with a small sample size. However, a bigger sample size would have provided a more thorough analysis of the end-users' needs and requirements, and would have permitted to generalize the results contained in this thesis in a better way. So we advise to work on this particular part of the project.

Another limitation is the small amount of use cases that have been scrutinized in section 3.1. We made a complete list of the different use cases available for this platform in section 3.3 but we analyzed just 9 of them (we analyzed the most important ones following the needs of the stakeholders). So future research on more use cases will be important when the platform will grow. Also, more research needs to be done on the possible use cases to implement in the platform to increase the list on section 3.3.

More existing technological solutions and current products and services on the market should be analyzed. A deep analysis of them will give us a better understanding of our competitors' products and the technologies that can be used in the SA of our platform.

Another limitation of this thesis is the superficial analysis made of single data sources that will be used to retrieve data needed for the use cases. We do not know completely the state and the accessibility of the data that is needed for these use cases so more research in this field would be needed. However, we implemented in our SA the "data conversion" service that will be able to convert data imported from many types of data sources into linked data. This SA permits us to manage different types of data sources and this smooth the poor analysis that we made for the available data sources. In addition, more data sources should be analyzed in order to be able to develop more use cases.

The SA and the UI of the platform should be tested in the real world when the platform is made. This is because we did not test them in the real world but we just tested the mockups with two professional user interface graphic designers and with five possible end-users.

Our solution to the trust issue should also be tested when the platform will be realized to know if the end-users are willing to trust us and to share their private information with us.

9.3. Recommendations for future work and research

Taking into consideration the mentioned research limitations and the overall project, recommendations for future work and research can be made.

The first recommendation is to better analyze the needs and requirements of the end-users. For instance by using a bigger sample of end-users to interview. This makes it possible to generalize the findings that we derived from a small number of interviews into more broadly tested findings.

The second recommendation is to look deeper into existing technological solutions and current products and services on the market. Thus it is possible to have a better SA with the latest technology available and it will give us better understanding of our competitors.

The third recommendation is to create more use cases when the platform will be implemented and will start growing. When the use base will start reaching a good base number, more use cases will need to be implemented. This is important because end-users will constantly need new use cases in the platform to take the best advantage from it. These new use cases can be found by finding data sources to connect to the platform and then by creating use cases that will use them.

The fourth recommendation is to analyze the different data sources that will be connected to the platform. This step will be important when the platform will be developed and when the first release of it will be made.

The next step of this project will be to develop the platform. We will use the SA and the UI explained in this thesis. When the code of the platform will be written, an analysis of the different parts of the SA and of the UI will need to be made. The different REST calls of the different SOA's services will need to be made and the different classes of every part of the SA need to be developed. The end-user interface will need to be converted into an HTML and CSS interface to permit the different end-user's devices to use it. The development of the platform will be the next main challenge because it will be both time and money consuming. However, it is a mandatory stage in publishing the first release of the platform.

Finally, when the platform is developed, our solution of the trust issue can be tested and so we will be able to find if the end-users will share their private information with our platform.

10. References

Apache Jena Fuseki. Retrieved 09/05/2015, from http://jena.apache.org/documentation/fuseki2/

ARQ - A SPARQL Processor for Jena. Retrieved 12/04/2015, from https://jena.apache.org/documentation/query/

Berners-Lee, T. (2006). Linked Data. Retrieved 10/04/2015, from http://www.w3.org/DesignIssues/LinkedData.html

Bizer, C., Heath, T., & Tim Berners-Lee, T. (2009). *Linked Data-the story so far.*

Board Service Oriented Architecture. Retrieved 11/05/2015, from http://help.board.com/Chapter_2_-_Board_Architecture_files/The_Client_Master_Server_Architecture.htm

Burger, A., & Dekker, P. (2001). The nonprofit sector in the Netherlands.

Cockburn, A. (2002). Use cases, ten years later. *Software Testing and Quality Engineering Magazine* (pp. 37-40).

Cross-Platform Authentication with Google+ Sign-In. Retrieved 11/05/2015, from http://www.slideshare.net/peterfriese/crossplatform-authentication-with-google-signin

Duval, A., & Brasse, V. (2014). How to ensure the economic viability of an open data platform.

Education - Campus Management System. Retrieved 11/05/2015, from http://www.comm-it.in/form/education.aspx

Francken, N. (2013). The Dutch foundation. Retrieved 03/05/2015, from http://www.step.org/dutch-foundation

GNU GENERAL PUBLIC LICENSE. (2007). Retrieved 26/04/2014, from https://www.gnu.org/copyleft/gpl.html

Heath, T., & Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology, 1*(1), 1-136.

Hill, D. J., Liu, Y., Marini, L., Kopper, R., Rodriguez, A., Futrelle, J., & McLaren, T. (2011). A virtual sensor system for user-generated, real-time environmental data products *Environmental Modelling & Software* (pp. 1710-1724).

Krafzig, D., Banke, K., & Slama, D. (2005). *Enterprise SOA: service-oriented architecture best practices*: Prentice Hall Professional.

LinkedIn - A Professional Network built with Java Technologies and Agile Practices. Retrieved 10/05/2015, from http://www.slideshare.net/linkedin/linkedins-communication-architecture

Mikoluk, K. (2014). MySQL vs PostgreSQL: Why MySQL Is Superior To PostgreSQL. Retrieved 09/05/2015, from https://blog.udemy.com/mysql-vs-postgresql/

Morgan, L., & Finnegan, P. (2014). Beyond free software: An exploration of the business value of strategic open source. *The Journal of Strategic Information Systems, 23*(3), 226-238.

Nechifor, A., Albu, M., Hair, R., & Terzija, V. (2015). A flexible platform for synchronized measurements, data aggregation and information retrieval *Electric Power Systems Research* (pp. 20-31).

Omitola, T., Koumenides, C. L., Popov, I. O., Yang, Y., Salvadores, M., Correndo, G., . . . Shadbolt, N. (2010). Integrating public datasets using linked data: challenges and design principles.

Santos, C., Kuk, G., Kon, F., & Pearson, J. (2013). The attraction of contributors in free and open source software projects. *The Journal of Strategic Information Systems, 22*(1), 26-45. doi: http://dx.doi.org/10.1016/j.jsis.2012.07.004

SPARQL 1.1 Overview. Retrieved 22/04/2015, from http://www.w3.org/TR/sparql11-overview/

SPARQL 1.1 Query Language. Retrieved 22/04/2015, from http://www.w3.org/TR/sparql11-query/

Stevens, M. (2002). The Benefits of a Service-Oriented Architecture. Retrieved 11/05/2015

Tekinerdoğan, B., & Akşit, M. (2002). *Synthesis-based software architecture design*: Springer.

Text searches with SPARQL. Retrieved 12/04/2015, from https://jena.apache.org/documentation/query/text-query.html

Van Gorp, P., Comuzzi, M., Jahnen, A., Kaymak, U., & Middleton, B. (2014). An open platform for personal health record apps with platform-level privacy protection. *Computers in biology and medicine, 51*, 14-23.

Van Oortmarssen, A., De Vries, M., & Van Loenen, B. (2015). Privacy op zijn plaats: Tussen willen weten en wetten: Geonovum.

What Is an Enterprise Service Bus? Retrieved 11/05/2015, from https://msdn.microsoft.com/en-us/biztalk/dd876606.aspx

What is Service Oriented Architecture? Retrieved 11/05/2015, from http://www.pnmsoft.com/resources/bpm-tutorial/soa-tutorial/

Wieringa, R. J. (2014). *Design science methodology for information systems and software engineering*: Springer.

Wilkes, D. S. a. L., & Forum, C. (2004). Understanding Service-Oriented Architecture. Retrieved 11/05/2015, from https://msdn.microsoft.com/en-us/library/aa480021.aspx

Appendix A:end-user interviews

We interviewed potential end-users in two interview rounds. We used the survey method to know the needs and the requirements of the end-users because "surveys are useful for implementation evaluation and problem investigation because they can provide information about real-world phenomena" (Wieringa, 2014). The disadvantages of surveys are that they can disturb the phenomena investigated and that they can investigate only a few aspects of the phenomena (Wieringa, 2014).

First round

In the first interview round we asked general questions that are summarized in Table 11: First round questions. We asked four potential end-users to answer these questions.

The details of each person that we interviewed are shown in Table 10. These end-users are representative of our end-users because we assume that every person will be able to use our platform. The sex of the person is not relevant because both females and males will be able to use the platform. Age is relevant because old end-users will have more trouble to use a web-platform, so the end-users that we interviewed are less than 50 years old. Whether they have children is not important because both end-users with and without children will be able to use our platform. Sex and the fact of having children are mentioned in Table 10 because we collected this information during the interviews but they are not relevant for this research.

Person #	**Sex**	**Age**	**Children**
1	Female	20-30	No
2	Female	30-40	Yes
3	Male	40-50	Yes
4	Female	30-40	yes

Table 10: Details of end-users interviewed in the first round

In Table 11 are the questions that we asked in the first round. For each question the person was asked to refer to the different contexts included in Table 12.We asked these questions because we wanted to know what end-users do when something is happening, what they are going to do when something is happening, what is important for them, how they know that something is happening, how they will solve the problem and if they are going to share relevant information about the situation with external parties.

Question #	**Question description**
1	How do you know that it is happening?
2	Which actions are you going to take?
3	What is the most important thing for you?
4	Who has to know that it is happening?
5	How to solve this problem?
6	Are you willing to share information with trusted end-users?

Table 11: First round questions

In Table 12 the different contexts that the person should have in mind when replying to the different questions are shown.

Context #	Context description
1	Fire in my home.
2	Theft in my home.
3	I have broken a body part.

Table 12: First round contexts

In the next table the replies that were given by the interviewees are shown.. If replies were similar they were merged into one single sentence in the table. A blank cell means that no reply for that particular combination of question and context was received.

Questions/context	Fire in my home	Theft in my home	I have a part of the body broken
How do you know that it is happening?	• From neighbors • Alarm signal and police calling	• Alarm signal	N/A
Which actions are you going to take?	• Call parents • Go home • Alert neighbors	• Call police, parents and neighbors	• Call for help of friends, parents and neighbors
What is the most important thing for you?	• Children, humans, photos, things that cannot be replaced • Mobile telephone for data • Documents and things with emotional value		
Who has to know that it is happening?	• Parents and friends • Firefighters	• Police, parents and neighbors	• Friends, parents and neighbors
How to solve this problem?	• Call the insurance company for compensation • Wait • If the fire is small try to stop it, if the fire is big do nothing	• Call the police, parents and neighbors	

Are you willing to share information with trusted end-users?	• Yes for a map of the house and where the gas and explosive materials are stored in the house • Yes for maps and all the information that can be helpful to prevent and extinguish the fire. • Not for sharing general information but in case of a fire it is ok to share relevant information but just by selecting the exact end-users that will receive this information	• Yes to share information with neighbors about a robbery that is going on in the house • Not for sharing general information. In case of a robbery it is ok to share relevant information but just by selecting the exact end-users that will receive this information	• Yes to share information about my health problems with the neighbors to get help • We have a good relationship with neighbors so we do not need a platform to contact them

Table 13: First round interview results

These results were helpful to get a general view of what end-users do in a particular context. We found that end-users normally know if there is a robbery or a fire in their house by the alarm signal, from the neighbors or from the police.

We found that end-users, when they know that there is a fire or robbery in their home, call their parents, police and neighbors and then they go home. This information can be important for the alarm sharing use case explained in section **Errore. L'origine riferimento non è stata trovata.**. If a person has a part of his body broken and cannot move outside the house, he will call his friends, family and neighbors for help.

People think that, when there is a fire in their home, the most important things for them are their children, any humans in the house, photos and things that cannot be replaced, data from mobile telephones, documents and things with emotional value.

They are willing to share their information, in case of a fire or a robbery, with their parents, friends, neighbors, firefighters and the police. In case they have a broken body part they are willing to share this information with friends, parents and neighbors.

When there is a fire at home, a person will call the insurance for get reimbursed, he will wait or, if the fire is small he will try to stop it. If a robbery is happening in the person's home, he will call the police, his parents or his neighbors.

To prevent fires in the house and to increase the probability of extinguishing the fire in a short time, end-users are willing to share information about the house with specific end-users (this information includes a map of the

house and the position of the gas stove). In case of a robbery, end-users are willing to share this information by selecting exactly the end-users that will receive this information. In case of a broken body part, a person is willing to share the information about his situation with his neighbors. However, they replied that they are already receiving help from their neighbors if they are in this situation so they do not need a web-platform to contact their neighbors.

Second round

In the second round we interviewed more end-users and we asked more specific questions to them. These questions were about the information that they would want to get from a web-portal, if they perceive the division of information over many different websites as annoying and if they would like to have any other functions in the platform.

The details of the interviewed end-users are summarized in Table 14. A blank cell means that the information was not provided. The end-users were both female and male and they had a wide age range. We decided to interview older end-users as well, because we wanted to include them in the end-users of the platform. In this second round we did not ask if people have children because that is irrelevant for this platform.

Person #	**Sex**	**Age**
1		70-80
2	Male	50
3	Female	40
4	Male	90
5	Female	40
6	Male	35
7	Female	45
8	Female	50
9	Male	65
10	Female	40
11	Female	35-40
12	Female	50
13	Female	40-50
14	Female	70
15	Female	30-40

Table 14: Details of interviewed end-users in the second round

The questions that we asked are listed in the next table. We asked questions that were divided into three categories. The first category is the "surrounding" category that contains all the information and services related to the surroundings of the end-user. The second category is the "house" category that includes all the information about and services for the house of end-users. The last category is the "energy" category that includes all the services and information related to energy and collective purchasing of things with neighbors. The aim of these questions was to know more about end-users' needs and how to satisfy them.

Question #	Category	Question
1	Surrounding	Do you want to know the level of traffic congestion in your surroundings?
2	Surrounding	Right now, do you not know about road construction or maintenance in the surrounding in time ?
3	Surrounding	Do you want to know if your neighbors ask for a building permit?
4	Surrounding	Do you want to know about a power failure in your surroundings?
5	Surrounding	Do you think that the alternatives in case of power failure are difficult to find?
6	Surrounding	Are you willing to share your belongings with your neighbors?
7	Surrounding	Are you familiar with an appliance sharing website?
8	Surrounding	Are you willing to lend belongings if your neighbors ask for it?
9	Surrounding	Do you know what is happening in your area?
10	Surrounding	Would you like to know what is happening in your area?
11	Surrounding	Do you know when there are local events in your area?
12	Surrounding	If you knew when there were local events in your area, would you go to them more often?
13	House	Do you perceive that the information about your house is scattered over a lot of different websites?
14	House	Do you perceive this scattering as annoying?
15	House	Would you share information about your house with other end-users?
16	House	Do you know how much your house is worth?
17	House	Do you want to know how much your house is worth several times each year?
18	House	Would you like to share pictures of the interior of your house with other end-users?
19	House	Would you like to see old pictures of the surroundings of your house in the platform?
20	Energy	Do you know how your energy usage is compared to your neighbors?
21	Energy	Do you know which devices in your house are consuming more energy?
22	Energy	If you knew which devices are consuming more energy, would you be able to save energy?
23	Energy	Would you like to know which energy saving options are available?
24	Energy	Would you like to know about collective purchases?

Table 15: Second round interview questions

The results are summarized in the next table, and are summarized in a binary way. We analyzed the information by labeling each answer as a yes/no binary answer. Then we summed them and derived the conclusion. The conclusion is yes or no if the answers were more than 70% in favor of one of these answers. In the other cases, the conclusion is neutral.

Question #	Yes	No	Conclusion
1	7	6	Neutral
2	5	7	Neutral
3	7	6	Neutral
4	10	2	Yes
5	7	3	Yes
6	8	2	Yes
7	0	9	No
8	8	4	Yes
9	5	8	No
10	8	2	Yes
11	4	6	No
12	7	3	Yes
13	5	1	Yes
14	4	2	Yes
15	0	3	No
16	2	0	Yes
17	1	2	Neutral
18	1	2	Neutral
19	1	3	No
20	0	1	Neutral
21	1	2	Neutral
22	1	3	No
23	1	0	Yes
24	2	2	Neutral

Table 16: Second round results

People seem neutral about the traffic congestion information, road construction information and about their neighbors building permits. Based on the answers that were received it can be concluded that end-users already know if there is road maintenance in their street because they receive a letter from the government with details of the construction work...

People think that alternatives, in case of a power failure, are difficult to find. End-users would like to know when the electricity will come back after a power failure.

End-users are willing to share belongings with neighbors, but they do not know the existing platforms that offer this service. If they knew these websites they would use them.

People do not know about local events of their area and, if they knew about them, they would join them more often.

People perceive that their information is scattered over different websites and this scattering is perceived as annoying.

People are not willing to share their information with other end-users. However, during the interviews, most of them said that they would share their information if it is not important information or if sharing it is really important, for example during a robbery. They are also willing to share their information just with specific end-users and with specific third parties.

People seem to know how much their house is worth and they seem neutral about getting information about the value of their house several times each year.

People are neutral about sharing of pictures of the interior of their house with external end-users.

People do not care about seeing old pictures of their house and of their surrounding in the platform.

People are neutral about information about the energy consumption of their neighbors and a comparison with their energy usage. People are also neutral about knowing which devices in their house are consuming more energy and they do not think that, with this information, they will be able to save energy. People would like to know which energy saving options are available for them.

Finally, end-users are neutral about collective purchasing of goods with neighbors.

Using these results we can determine the necessary requirements for the platform. These requirements are discussed in section 2.1.3 where the complete list of requirements is provided.

Appendix B: data sources that will not be used for the software architecture in this platform but that can be used for future use cases.

The following data sources were found using several methods. First, we searched for data sources in data source collection websites such as "data.overheid.nl" and "nationaalgeoregister.nl". These websites collect a huge amount of public datasets that can be used in our platform.
Secondly, we found data sources through partnerships and external parties that will develop use cases for the platform.
Finally, we other data sources were found using web search engines such as Google.

ID	Name	Description	How to retrieve it
1	Traffic density	The traffic density of the street where the end-user lives	
2	Perceived safety	Perceived safety of end-users divided per postcode	https://data.overheid.nl/data/dataset/onveiligheid-buurtkenmerken-2008-2009
3	Quality of living	Quality of life in the suburb	https://data.overheid.nl/data/dataset/leefbaarometer-1998-2010
4	Weather forecast	Information about the weather forecast with added information connected to it. For instance: - garden info system: informing you about actions needed, watering based on weather, soil, ground water level, season, etc. - storm warning: close your windows - heat levels: expected warm days in which rooms on the top floor will be too warm to sleep conveniently. This added information will be processed by our platform.	
5	Temperature inside the house	Temperature data from the temperature sensors placed in different parts of the house	

6	House tax details	Amount of taxes to pay for the house.	
7	Information for an activity in the city	End-users will receive an alert if there is an event near their postcode. The alert can be for traffic or for high sound levels. This can also be combined by showing the details of the activities that are planned in the city.	
8	Monuments	If the house is a monument there will be information about its history or for instance renovation restrictions	
9	Information about the street	History of the street, important events, important end-users that lived there	
10	Cables and pipes in the underground	Information about the pipes (water and gas) and cables (electricity and telephone) that are underground near the house	
11	Local emergency telephone numbers	Local police telephone number and the nearest hospital telephone number	
12	Nearest hydrant	Nearest hydrant of the end-user	
13	Map of the surroundings	A map of the surroundings of the end-user	
14	Local news	Local news of the city and the surroundings	
15	Parks information	The nearest park and its details	
16	New photos and videos of the surroundings	New public photos and videos of the surroundings	
17	Pollution statistics and forecast	End-users will be able to see pollution statistics and forecast of their area and advice about	

		how to protect themselves from the pollution.	
18	Parking places	Nearest public parking places and their hourly fees	
19	Average temperature, rainfall and climate	Data about the average climate of the area	
20	Energy class of the house	Energy class of the house	

Table 17: Data sources that will not be used for the software architecture of this platform but that can be used for future use cases

Appendix C: mockups of the interface

This chapter has been left out because it is part of the private version of this thesis.