**Master Dissertation**

**EE5500**

Markus Just

Student number: 1644609

Dr. Alizrea Mousavi

Electronic and Computer Engineering

School of Engineering and Design 

Tuesday, February 20, 2018

Table of contents

[Introduction 4](#_Toc506900428)

[Background to the project 6](#_Toc506900429)

[Regulations 11](#_Toc506900430)

[Drinking water quality 11](#_Toc506900431)

[External factors 12](#_Toc506900432)

[Aims and Objectives 15](#_Toc506900433)

[Experimental/investigative methods to be adopted 18](#_Toc506900434)

[Theoretical concepts 19](#_Toc506900435)

[Data dissemination 19](#_Toc506900436)

[Data integration 20](#_Toc506900437)

[Generic cloud-based service provider 20](#_Toc506900438)

[Context view 20](#_Toc506900439)

[Admin 21](#_Toc506900440)

[User 21](#_Toc506900441)

[DataAggregator 21](#_Toc506900442)

[External data source 21](#_Toc506900443)

[Use cases 22](#_Toc506900444)

[Use-case “Manage Users/Roles” 24](#_Toc506900445)

[Use-case “Configure Views” 25](#_Toc506900446)

[Use-case “Viewing information” 25](#_Toc506900447)

[Domain Model 25](#_Toc506900448)

[Use-case “Manager users/roles” 26](#_Toc506900449)

[Use-case “Viewing information” 26](#_Toc506900450)

[Use-case “Configure views” 27](#_Toc506900451)

[Implementation 28](#_Toc506900452)

[Technology overview 28](#_Toc506900453)

[Cloud-Solution 28](#_Toc506900454)

[Programming Language and Frameworks 28](#_Toc506900455)

[Software Architecture 29](#_Toc506900456)

[User management 29](#_Toc506900457)

[Data analysis 29](#_Toc506900458)

[Project plan 30](#_Toc506900459)

[Time-plan 30](#_Toc506900460)

[Deliverables or specific outcomes 33](#_Toc506900461)

[References 34](#_Toc506900462)

# Introduction

Today it is natural that clean water is available whenever it is needed. Water is not only used for drinking, it is also used for cleaning, cooking, washing or many other things in daily life. Besides for individuals it is important for whole industries like agriculture industry, catering industry, food industry and many more. It is one of the most important basics in today’s life that clean water is accessible for everyone with a certain quality and quantity. To make sure that clean water is present all the time, there are industrial processes, which use water and wastewater treatment plants to make use of natural water resources and sewage to produce clean water which can be used by consumers. The following picture shows the process for natural water resources:



Figure 1 - water treatment [1]

The purpose of this report is to give an initial overview about the master dissertation project, which topic is water resource planing in the UK with the help of modern web technologies.

This project is strongly realted to another project, so this report will briefly describe the general idea and will then focus on the part which is relevant for this dissertation. After the information about the backgound of the project, there will be a part which describes the initial survey done. Furthermore, this report gives information about the aims and objectives of the dissertation and the expected outcomes as well as the time plan for the project.

## Background to the project

“Customers’ top priority for water services is a safe, reliable supply of water at a price they can afford” [2]. Because of different factors the risk of droughts in the UK increases steadily and there have also been some droughts during the last 40 years which were worse than those that had been used as the basis of planning. [2] One factor which is responsible for the increased risk of a drought is the climate change as well as the resulting increased evaporation during the warm months of the year. Another factor is the population growth, which is estimated to be between 6,6 million and 16 million by 2040 for England and Wales. Especially regions with the least resilience regarding water resources are subject to most population growth and climate change. [2]

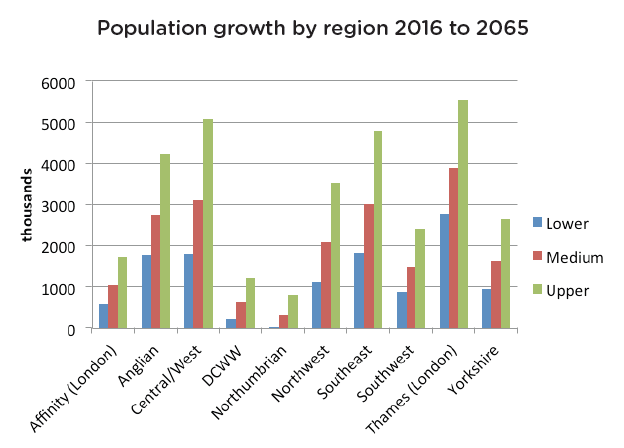


Figure 2 – Population Growth [2]

To protect customers as well as the industry from the consequences of serious droughts and to enable an economical and environmental efficient use of water resources, politics and the water industry have to make strategical long-term and short-term decisions regarding water infrastructure to be able to gain resilience.

The problem is that in water industry there is no technology which provides information about the overall state of the current water resources, neither for consumers nor for providers or political decision-makers. Because of this missing knowledge, there is a lack of efficiency regarding different aspects, e.g. energy consumption or cost. It is not possible to achieve an efficient water resource planning without information about important key indicators and predictions.

The basis of this project is the data acquired by modern sensors of water and wastewater treatment plants. This data is categorized in the industry’s key performance indicators. Those indicators are Carbon Footprint, Energy Consumption & Generation, the yield of the By-Products and the Productivity, which means the performance of the Equipment and Reliability. It has to be dealt with the real-time data as well as with historical data. As already mentioned, the acquisition, harmonization and provision of the data from different water providers is done in another project. This dissertation focuses on the data analysis and data representation. That’s why it has to be analyzed which stakeholders there are to create a concept how the data analysis has to be done and which factors has to be considered. A source for this kind of knowledge is Water UK:

“Water UK is a membership organisation which represents and works with the major water and wastewater service providers in England, Scotland, Wales and Northern Ireland”. [3]

Water UK provides all kind of information about water supply, wastewater and regulations regarding those processes. Focusing on the regulators, Water UK makes the following statement:

“A wide range of organisations work with water companies to ensure customers get the best services for the best possible price, and that the environment is protected.” [4] The mentioned organisations are categorized in governmental organisations, regulators, consumer watchdogs and water companies. These categories will be relevant for the concept of the data visualization:

*Governmental organisations:*

[**Defra**](https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs) **(**[Department for Environment, Food & Rural Affairs](https://www.gov.uk/government/organisations/department-for-environment-food-rural-affairs), [Environment Agency](https://www.gov.uk/government/organisations/environment-agency)):

UK government department responsible for looking after natural environment. Defra sets the overall rules for water services in England and is responsible for water quality and orders regarding droughts.

[**Welsh Government**](http://gov.wales/?lang=en): devolved Government for Wales which works to help improve the lives of people in Wales and make the nation a better place in which to live and work. The Welsh Government sets the overall rules for water services in Wales. [4]

*Regulators:*

[**Drinking Water Inspectorate**](http://www.dwi.gov.uk/): the DWI was formed in 1990 and provides independent reassurance that water supplies in England and Wales are safe and drinking water quality is acceptable to consumers. [5]

[**Environment Agency**](https://www.gov.uk/government/organisations/environment-agency): non-departmental public body which is responsible for regulating industry waste, as well as water quality and resources in England. They are also responsible for managing the risk of flooding from rivers, reservoirs, estuaries and the sea. [4]

[**Natural England**](https://www.gov.uk/government/organisations/natural-england): the government's adviser for the natural environment in England, helping to protect England's nature and landscapes for people to enjoy and for the services they provide.

[**Natural Resources Wales**](https://naturalresources.wales/?lang=en): works to ensure that the environment and natural resources of Wales are sustainably maintained, enhanced and used, both now and into the future.

[**Ofwat**](https://www.ofwat.gov.uk/)**:** the economic regulator for the water and sewerage sectors in England and Wales. It works in the interest of customers by setting price limits, ensuring companies run efficiently and encouraging resilience. [4]

*Consumer Watchdog:*

[**CCWater**](http://www.ccwater.org.uk/): The Consumer Council for Water promotes consumers' interests to governments, regulators and water companies. They also provide a free advice and complaint handling service for consumers, research their views on key topics, and keep them informed on the issues that affect their services. [4]

All these stakeholders have different kind of interests in the data, which has to be considered when doing a data analysis with a visualized result. What exactly these interests are must be researched beforehand which is one of the first steps during the project.

## Regulations

### Drinking water quality

In the UK, drinking water quality has the highest priority for water suppliers. Independent drinking water inspectorates regularly[[1]](#footnote-1) check and ensure the quality of the water provided to customers. The national regulations specify strict standards derived from the EU Drinking Water Directive, which is based on advice from the World Health Organization (WHO). [6]

It is the requirement of each member state of the EU to translate requirements of the EU Drinking Water Directive to local laws. The UK follows this requirement and adds also additional rules, which leads to very high standards in regards of drinking water quality. As mentioned before, the EU and the UK regulations are based on the advice of the WHO, which are regularly updated mainly because of new gained knowledge. To make sure that the regulations of the EU and the UK are updated accordingly, the European Commission review the current stadards at least every five years and update them if it is necessary. [6]

Besides the drinking water quality the EU directive lays down strict requirements for monitoring, analysing and reporting of measured data as well as requirements about actions which has to be taken if standards are exceeded. [6] The DWI states that water companies have to ensure that samples for e.coli, coliform bacteria, colony counts, residual disinfectant, turbidity and nitrite are taken at a specific frequency from the point the water leaves a treatment works. [7]

## External factors

The data analysis which has to be done is based on two data sources. On the one hand it is the data which is collected on the plant level. This means real-time and historical data about the productivity, the energy consumption or the water quality measured by the sensors of the water treatment plants. On the other hand there are external data sources which will be required to perform fitting analyses. So it is important to identify which external factors are relevant for the water resource planning and what’s the impact of these factors.

Maybe the most important external factor is the weather situation. It has impact on the total availablity of water, on the demand for water by the people and on the water quality. The Met Office[[2]](#footnote-2) states [7]: “With climate change predicted to increase, the likelihood of significant weather events from extreme rainfall to heatwaves, will have a significant impact on the water industry.“ Extreme weather conditions may have a serious impact on water quality e.g. too much rain can cause sewers to flood and overflow into water course. Heatwaves with few rain may lead to droughts which could result in a reduction of availability of water as well as a higher demand of the consumer

s. If less water is available then the overall cost of delivering water to the end consumer may increase for the industry. With the knowledge of upcoming extreme weather conditions, actions could be taken early by regulators or consumers.

As mentioned before the cost of the whole process is also interesting for the stakeholders. Ofwat sets price limits for customers, so that the providers have to work efficiently and look after resilience to avoid sudden increase of cost. The following image shows how the price for water is set by a specific water provider (south east water):

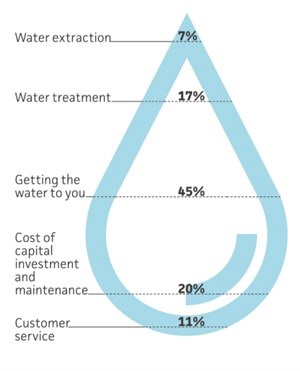


Figure 3 - how the price is set [8]

Almost 20% of the price is related to the water treatment process. So the absolute cost of the treatment process will influence the overall cost for whole process. This information could be very useful for water companies and may lead to strategic decisions depending on the current situation. From 2016 to 2017 the average household water and sewage bill in England and Wales was about £389. [9] The prices increased by £2, which is part of a 5-year-plan confirmed by Ofwat. This plan enables companies to invest £44 billion over 5 years in better services, greater resilience and environmental improvements. [9]

# Aims and Objectives

The main objective of this project is to develop a generic cloud service which is able to collect, analyze and visualize data independent of its origin and meaning. The previously described problem will be taken as an example use case and should show the validity of the concept as well as possible problems.

Taking the WWTP example it should be possible to integrate data from different sources, analyse them and create different results and representations, which can be categorized in different levels:

- strategic level: level which enables users to create strategies (e.g. price shouldn’t go over x)

- operational level: engineering level combined with external factors (e.g. weather)

- engineering level: lowest level/plant level

For example an end consumer wants to see different kind of data than a representative of a water provider company or a politican which aims towards developing a strategy for water resource management. This should be achieved by analyzing real-time data and/or historical data combined with different external factors like weather conditions. It is necessary to have a fitting security concept to guarantee that every user role can only see information which it should be allowed to see. For example a member of a water provider company shouldn’t be able to see sensitive data from another provider company (of course a consumer shouldn’t also be able to see sensitive copmpany data).



Figure 4 - system overview [9]

The illustration shows that the water and wastewater treatment plants of different water providers are equiped with different kinds of modern sensors, measuring data regarding several key indicators like Carbon Footprint, Energy Consumption & Generation, the yield of the By-Products of the processes, and the Overall Performance of the Equipment and Reliability (Productivity). This sensor data from different plants is collected and will be migrated to a common data model.

Each service provider categorizes the data from its plants in different indicators. The Large Network Performance Collider collects this data and converts it into a common model. This collected data is the basic information for the data analysis which has to be done to create the representation for stakeholders to enable strategic decision making.

The KPIs which are acquired at plant level and communicated over the service providers’ internal networks will be translated to a provider-specific evaluation of three aspects:

*Sustainablity:* Efficiency, water networks, waste

*Economic:* Economic factors (cost)

*Societal:* Well-being of public

Those results will be collected from different service providers to create an overall view which enables strategic decision-making. Real-Time data should be combined with historical data to create a systematic knowledge and data engineering capability. Additionally external factors have to be considered as well at this point as described in chapter (REFERENZ).

# Experimental/investigative methods to be adopted

It has to be analyzed which cloud technologies will be used to create a fitting representation of the current situation considering external factors and/or historical data. Furthermore a strategy has to be developed how to analyze the harmonized data from the different plants and create meaningful results, depending on the users’ request. It must also be evaluated which external data sources must be used to produce useful information. Another important task will be the role concept for the application. A fitting security strategy has to be developed to ensure that every user is only allowed to see what he/she should see. Critical data has to be secured from unauthorized access to avoid the misuse of information.

Due to the fact, another project is working on the problem of the data aqcuisition and data harmonization of different data sources (different providers), there is a need for “simulated” data which can easily replaced by the real data when it is available.

For the visualization of the results, a fitting user interface must be developed. It has to be researched which user will be interesed in which kind of information and after that the user interface has to be designed by using Mock-ups.

# Theoretical concepts

## Data dissemination

The concept of data dissemination is present all over the Internet. Every second millions of internet users receive data from different servers all over the world. This is made possible through communication protocols such as HTTP (Hyper Text Transfer Protocol) in combination with Internet Browsers and HTML (Hyper Text Markup Language). Using this technologies a user is able to navigate through the world wide web and view websites presenting data to the user. Another widely spread mechanism for data dissemination thorugh the internet is E-Mail communication. Using the SMTP protocol (Simple Mail Transfer Protocol) E-Mail is used to disseminate data through the internet or through intranet systems which are used by companies to enable their staff to communicate. [11]

Despite the efficiency of electronic means of data dissemination, there are still drawbacks which may take a long time to overcome, if at all. Privacy is one of the most common problems with electronic data dissemination. The internet has thousands of loop holes where people can peep into the private lives of other people. Security is also a related problem with electronic data dissemination. Every year, millions of dollars are lost to electronic theft and fraud. Every time a solution is found for a security problem, another malicious programs spring up somewhere in the globe.

Many companies set up precautionary measures against security invasion in their information systems. Some set up user accounts with varying privileges to data access. Many set up internet firewalls and anti virus software on their computers to prevent intrusions.

Data dissemination is a very substantial aspect of business operation. Most of today’s businesses are data driven. It is a common scenario where business organizations invest millions for data warehouses including hardware, software and manpower costs, to make data dissemination fast, accurate and timely. Information gathered from disseminated data form as basis for spotting industry trends and patterns and decision making in companies. [11]

## Data integration

Data integration is the combination of technical and business processes used to combine data from disparate sources into meaningful and valuable information [12]

# Generic cloud-based service provider

## Context view

The context view of a system describes the system under development (SuD) as a blackbox and how it is embedded in its environment. This view is used early in the development process to define the scope of a project and it can be considered as a connections between the written requirements and the system architecture. This view aims towards clarifying the system boundaries and the interfaces to the outside world.

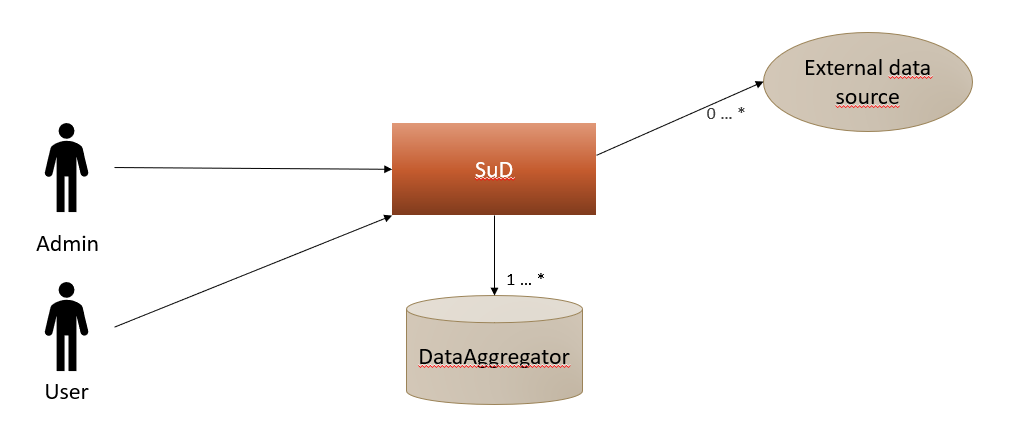


Figure 5 - context view

### Admin

### User

### DataAggregator

### External data source

## Use cases

Das Ziel jedes Softwareentwicklungsprozesses ist es, eine Software zu entwickeln, die ganz bestimmte Anforderungen erfüllt. Die Entwicklung einer Software fängt mit der Zielsetzung an: Die Software soll, wenn fertiggestellt, die zu Beginn des Entwicklungsprozesses festgelegten Anforderungen erfüllen. Niemand entwickelt Software grundlos und wartet darauf, dass zufällig eine für irgendetwas brauchbare Software vor seinen Augen entsteht. Es gibt immer Ziele, die erreicht werden sollen.

Leider sind diese Ziele nicht immer klar definiert. Man nimmt sich zum Beispiel vor, einen Online-Shop zu entwickeln, und stellt dann während des Entwicklungsprozesses fest, dass es unendlich viele unterschiedliche Funktionen in einem Online-Shop geben kann und man sich eigentlich nie klar gemacht hat, was man denn nun für Funktionen im Detail braucht. Man muss den Entwicklungsprozess daher wiederholt unterbrechen und inne halten, um sich zu überlegen, welche Funktionen, die einem bei der Entwicklung gerade eingefallen sind, notwendig sind und welche nicht.

Besonders schwierig wird die Situation, wenn der Auftraggeber des Entwicklungsprozesses nicht gleichzeitig der Entwickler ist. In diesem Fall kann der Entwickler nicht entscheiden, welche Funktionen notwendig sind - dies weiß nur der Auftraggeber. Dies führt zu einem ständigen Frage-Antwort-Spiel zwischen Entwickler und Auftraggeber, wenn der Entwickler nicht - noch schlimmer - die Entscheidungen selbst trifft und hofft, dies jeweils im Sinne des Auftraggebers zu tun.

Wenn Anforderungen an die zu entwickelnde Software nicht zu Beginn des Entwicklungsprozesses klipp und klar sind, wird der Entwicklungsprozess an sich unnötig erschwert. Denn das, was Sie entwickeln, richtet sich nach den bekannten Anforderungen. Jede Anforderung, die Ihnen oder Ihrem Auftraggeber später einfällt, führt dazu, dass Sie das, was Sie bisher entwickelt haben, ändern müssen. Denn die neue Anforderung hatten Sie logischerweise in Ihrer bisherigen Entwicklung nicht berücksichtigt. Grundsätzlich gilt, dass je später Anforderungen in einem Entwicklungsprozess bekannt werden, umso aufwändiger und daher teurer der Entwicklungsprozess wird. Anders gesagt: Wenn alle Anforderungen von Anfang an bekannt sind, bevor der Entwicklungsprozess gestartet wird, wäre das ideal.



Master User

Service User

<<include>>

Figure 6 – Use Case Diagram

### Use-case “Manage Users/Roles”

The master user, who belongs to the service providing company is able to manage the user roles and the specific users. Managing these subjects means to configure the authorities and privileges regarding the data which will be presented within a view. This ensures that only privileged users are able to see sensitive data, e.g. an employee of a water provider company should be allowed to view sensor data from his company’s water plants, however a customer shouldn’t be allowed to view this.

### Use-case “Configure Views”

The master user is able to configure the individual view for a specific user or user role. A view consists of one or multiple GUI-components representing a data set. An example for a view is a table with real-time sensor data along with a diagram which visualizes the development of sensor data over the last 24 hours.

### Use-case “Viewing information”

The service user (the customer) is able to see the pre-configured views after he logs in to the application.

## Domain Model

The creation of data models is often one the first steps in software development to build the foundation of the software architecture. Data models are independent of the used technology and can be used for different target systems, e.g a desktop application, web service or a database. Data models are useful for developers to get an idea how to build the system, however it is also useful for customers to clarify the purpose of system together with the developers. In general, data models show objects of the real world and the static relations between these objects. In this project it is a bit different, because the system should be able to deal with data which is not known during development. So the data model must be build with generic data objects.

### Use-case “Manager users/roles”

A user of the system is assigned to a user role. Every role has a set of authorizations regarding specific data sets.

User

User Role

Set of authorizations

Data set

### Use-case “Viewing information”

From a high level view there aren’t many domain objects which are involved in the default use case. A user of the system is assigned to a user role which is defined by its permissions. The user sees the user view (consisting of GUI-components) which visualize data from one or multiple sources.

### Use-case “Configure views”

Regarding the data model, the use case “Configure views” doesn’t differ from the use case “Viewing information”, because the domain model makes no representation about the initiator of a use case or the dynamic behaviour.

User

User Role

User view

Set of authorizations

Data source

GUI-component

Figure 7 - initial data model

# Implementation

## Technology overview

* Java
* Portable
* Proven build systems
* Maven
* Easy setup

### Cloud-Solution

* High scalable
* Cost-efficient
* AWS

### Programming Language and Frameworks

* Flexible frontend structure
* Support of drag and drop operations
* Java
* Charts
* Data-Binding
* Event-based
* Vaadin

### Software Architecture

### User management

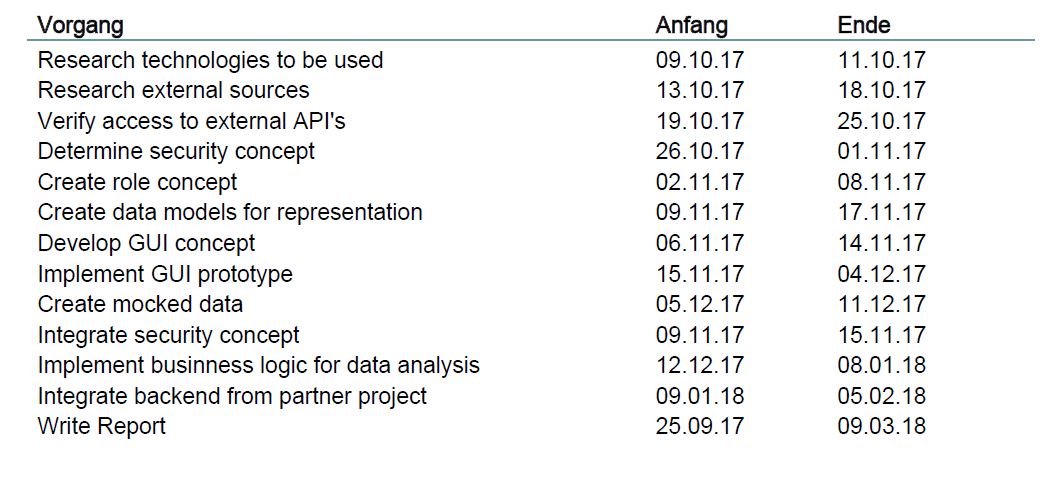
* Roles
* Users

### Data analysis

# Project plan

# Time-plan

The dissertation project should be finished until March 2018. To ensure that this project will be finished until then, an initial time plan is created which contains the key activities that have to be done with a time estimation. This time plan can be used as a tracking mechanism to detect possible delays early and take actions against it early (TODO: reformat time table):



Screenshot 8 - Time Plan

It is not guaranteed that all the tasks are found on creating the initial plan. It is likely that further tasks emerge during the research and the development. Nevertheless the initial time plan gives a good orientation which will help during the work process. The following section describes the tasks briefly to give a clearer overview and understanding over the project:

|  |  |
| --- | --- |
| Task | Description |
| Technology research | The initial research about the software technologies used to achieve the project aims. It is necessary to make decisions about several technologies like databases, cloud providers, frameworks etc.  Those choices have to be justified by showing the adventages over available alternatives. |
| External source research | There will be external sources which have to be accessed by the web app to enable analysis of different aspects. During this task the required external sources have to be identified. |
| API verification of external sources | After the identification, the access to the external sources has to be verified. It must be determined if these sources offer an external API and under which requirements they can be accessed. |
| Determine security concept | It is important that not every information is accessible to everybody, so a security concept is required to protect the web application. This concept has to be developed during this task, also considering provided services by the hosting provider. |
| Create role concept | There will be a role concept to distinct different users (e.g. admin user). The required roles have to be identified and the possibilities for the technical implementation must be verfied. |
| Create data models | A domain model has to be developed in this task. It should show the relations between the different entities and should be the basis for the presenetation of the data. |
| Develop GUI concept | Outcome of this task will be a first mockup of the web application GUI. Static views as well as possible user interaction should be part of this concept. |
| Implement GUI prototype | The outcome of the previous task should be implemented as a protoype with placeholders. |
| Create mock data | It is likely that it won’t be possible to access real data during the development, so it will be necessary to create mock data, which should be easy replaceable as soon as the integration of data providing systems is possible. |
| Integrate security concept | The defined security concept should be implemented in this task. |
| Implement business logic for data analysis | During this task the core of the web application should be implemented. The data from the different sources must be accessed and a data analysis will be performed which results in a meaningful visualization (HIER NOCH ETWAS DAZU SAGEN) |
| Integrate backend from other project | If possible, the data from the data aggregation and harmonization project should be accessed and used by the web application to perform the data analysis. |
| Write report | All the steps should be documented in this dissertation as well as the initial research. |

# Deliverables or specific outcomes

The expected outcome of this dissertation is a working prototype of a cloud application which visualizes data regarding water resource planing considering real-time data, external factors and historical data. This prototype should use a fitting security and role concept. The overview should help stakeholders to get an idea of the current state in their area of interest and make them able to make strategic decisions to achieve a certain change. Besides the representation a data analysis has to be done by using the acquired data of the partner dissertation. Along with the technical deliverable in form of a web application there should also be documentation about how the presented data is created and categorized and which indicators are used for evaluation.

# References

|  |  |
| --- | --- |
| [1] | „Water supply,“ [Online]. Available: https://www.water.org.uk/about-water-uk/water-supply. [Zugriff am 24 09 2017]. |
| [2] | Water UK, „Water resources long-term planning framework,“ Water UK, 2016. |
| [3] | „Water UK,“ [Online]. Available: https://www.water.org.uk/about-water-uk. [Zugriff am 23 09 2017]. |
| [4] | „Regulation,“ [Online]. Available: https://www.water.org.uk/about-water-uk/regulation. [Zugriff am 24 09 2017]. |
| [5] | „DWI,“ [Online]. Available: http://www.dwi.gov.uk/about/index.htm. [Zugriff am 23 09 2017]. |
| [6] | „Watrer UK: Drinking water quality,“ [Online]. Available: https://www.water.org.uk/policy/drinking-water-quality/water-quality-standards. [Zugriff am 03 01 2018]. |
| [7] | „Water,“ [Online]. Available: https://www.metoffice.gov.uk/services/industry/water. [Zugriff am 26 09 2017]. |
| [8] | „South East Water,“ [Online]. Available: http://www.southeastwater.co.uk/your-account/about-your-bill/water-charges/how-much-does-your-water-cost. [Zugriff am 29 09 2017]. |
| [9] | A. Mousavi und E. Katsou, „WWTP-Global-300617,“ 2017. |
| [10] | „Water UK,“ [Online]. Available: https://www.water.org.uk/news-water-uk/latest-news/household-water-and-sewerage-bills-2016-17. [Zugriff am 28 09 2017]. |

1. The checks take place each year [↑](#footnote-ref-1)
2. National meterologic service of the UK [↑](#footnote-ref-2)