UiT
THE ARCTIC UNIVERSITY OF NORWAY

Faculty of Science and Technology, Department of Computer Science

RoadAhead — Removing Uncertainty in Travel

Creating a Data Warehouse for Green Transportation Nudging

—

Håkon Wallann

INF-3981 Master's Thesis in Computer Science - June 2019

Abstract

This paper describes a data warehouse approach to environmentally friendly transportation nudging.

Transportation makes up a significant part of global carbon emissions. These emissions impacts both the climate and the health of individuals. As such, efforts should be done to address transportation patterns and habits. In addition to the reduction of air pollution, making people more active through active transportation has health benefits of its own.

Nudging is a tool meant to affect a person's choice in a non-coercive manner. An example of a transportation nudge is giving reminders of when a bus to a chosen destination is close to a nearby bus stop. The goal of this project is to nudge people to use healthier and greener transportation options by providing *certainty* in travel. We believe that we can affect a person's choice by providing them with relevant information about their travel paths.

In this thesis research into relevant data sources are investigated and an initial system was created to reach the goal of this project. The system created provides information relevant for in-city traveling based on data analysed from multiple sources. In addition to a few simple analyses implemented in the initial part of the system, more are discussed in this thesis.

An IoT device meant to measure snow height was developed and tested to be part of the first-party provided data. The design used a sonar sensor to measure the distance to the snow. Experiments done during this project shows flaws with using such a sensor during bad weather. Improvements to the device is also discussed in this thesis. In addition to the initial system, research into possible enhancements and the future of the system is investigated.

Acknowledgements

- I would like to thank Anders Andersen for his help and guidance in this project. His trust in this project and his research into the NUDGE project was of great help in the making of this thesis.

- Thank you to Pontus Edvard Aurdal for his help with *VisualBox*, both with the usage and creation of components. His work in visualization was vital for the implementation of this project.

- For help with sensors, 3D-printing and everything IoT related, I would like to thank Øystein Tveito. The integrity of the IoT device's hardware and casing was secured through his guidance.

- I would like to thank Fredrik Høisæther Rasch for his help with *GDAL* and map configurations. His knowledge of how to utilize the *GDAL* tool helped in visualizing data to users in a way that they could easily understand.

- I would also like to thank Jørgen Aarmo Lund for his help in proofreading this thesis. His pointers in structuring and grammar helped in increasing the quality of this thesis.

Contents

List of Figures

1

Introduction

Most people's daily routine involves traveling to and from work, education or activities. Part of this routine is preparing for travel, by looking at road and weather conditions. This could be done through something as simple as looking outdoor and checking the temperature, or by using a travel planner application like *Google Waze* [1]. In some cases, the amount of preparation done can decide if a day will be a good or a bad one. An example of this is preparing for rain on an overcast day.

The information that can help prepare a person is often freely available. This is especially true in developed countries where data collection is widespread. While information is often collected in abundance, using it effectively is not always easy. Finding accurate and useful sources can be difficult. After finding such sources, the data must be analysed. This is something that can be hard to do for people without specialized knowledge. Applications that can help in collecting and analysing data can have difficulty in recommending the best transport option. What is considered *best* might be subjective. The bicycle can only be the best option if one is available and the user knows how to use it. In this project, the best transportation method is one that takes in the wants and needs of the user and tries to nudge them to travel more environmentally friendly. This thesis will investigate what kind of information could be useful for this, and how to extract and analyse it.

1.1 Private Transportation

A private car is the preferred transportation choice for many people. According to *SSB*[1 2] [2], about 76,3% of the traveling in Norway in 2017 was done with a private vehicle. To get from A to B you must be outside to some extent, which can be unpleasant in bad weather. Cars, in addition to being able to transport people easily and quickly, protect us from weather conditions. This makes cars a *safe choice* when one wants to travel comfortably. A cars headlight provides visibility and spiked wheels ensure a firm grip on the ice. The enclosed space of the car helps with heating and protection from rain, snow, and wind. Research show that users of private vehicles gains more psychosocial benefits, like a feeling of mastery and autonomy, than people using public transportation [3]. In STEG [4], it was also found that fervent car drivers think that: "*... the car is a symbol of freedom and independence, a status symbol ...*" *[4]*

In recent years the knowledge about climate change and the impact of cars on carbon emissions has increased [5]. Even with the research into the development of environmentally friendly cars increasing, most cars still use fossil fuels. Of the about 2,75 million private vehicles registered in Norway in 2017 less than 400 thousand was registered as either an electric or a hybrid vehicle [2]. The great number of fossil fuel driven cars on the roads has a great impact on the environment. In addition, the environmental impact, the cost of fuel, at least in Norway, is on the rise. Over the last 10 years, the price of gasoline and diesel has risen by a bit more than 4 NOK per liter[3]. Rising fuel prices and an increasing awareness of the impact on the environment makes alternative transportation increasingly more attractive.

The widespread usage of cars in cities has a negative impact on the quality of life of those living there. While the option of travel directly to your destination is positive, there are several negatives too. Research into the health detriments brought by air pollution ranges from heart disease and lung cancer and asthma attacks [6]. According to *WHO*[4] around 7,6% of deaths in 2016 worldwide was caused by ambient air pollution [7]. The pollution of city air is one of the impacts cars have in heavily trafficked areas. In addition, traffic accidents happen every year. These types of accidents can damage both people and the environment. Information about the roads can decrease these accidents to some degree. An example of this is informing people about where icy roads are likely to be. Decreasing the total number of cars in traffic at the same time, like with carpooling, can also help.

1. Statistisk Sentralbyrå
2. `https://www.ssb.no/transport-og-reiseliv/faktaside/bil-og-transport#blokk-1`
3. `https://www.ssb.no/statbank/table/09654/chartViewLine/`
4. World Health Organization

A busy person might only be outside when traveling. This could also be the only time during the day that she is active. Keeping the body active and spending time outside, or at least in clean air, is healthy. Physical activity has a positive impact against multiple health problems, like cardiovascular diseases and cancer. Research into active travel shows that walking or biking to work has health benefits in all but the most extreme cases of air pollution [8, 9]. Preparation and information are necessary to travel comfortably without a private vehicle. This is because you must travel without the vehicle's protection, and you must prepare for longer commuting times.

Alternatives to private transportation usually have longer travel times and are more exposed to the weather. Examples of such alternative are walking or taking a bus.

1.2 Affecting Transportation Decisions

Which mode of transportation to use is decided by a cost-benefit analysis with subjective and changing variables. One day the factors impacting the decision could be the temperature and the distance to the destination, while other days *preference* could the only factor. The arbitrariness of the factors can be mitigated by looking for common ones that affect all humans. An example of this is that low temperature makes us feel cold.

A common part of cost-benefit analysis is that it requires a certain amount of knowledge to make a satisfactory decision. It is therefore possible to affect user's analyses by building a system that provides relevant information. A problem with creating such a system is the difficulty in knowing what type of information to provide. This is especially hard when the factors deciding how to travel is unknown and maybe changing over time. A scale of how much information to provide and its usefulness must be decided upon.

On the ends of the information provision scales are providing a decision and no information, and providing a lot of information but no decision. Both edges are flawed. Too little information requires the users to trust us explicitly, while too much information will only confuse them. In the end, a compromise is wanted for this project. This compromise presents analyses of collected data and gives advice and warnings based on those. Additional information could also be provided to those who want to understand the analyses in more detail.

1.3 Data Categories

We categorize collected data into **four different types** based on what the information is about: *the road*, *weather conditions*, *parking*, or *people*. We use these categories because they describe all parts that are connected to traveling.

- A person must travel on a road, or a path, to get somewhere.
- During the trip, she must deal with weather conditions in various ways, like using a raincoat in bad weather.
- A trip must start and stop somewhere, which falls under parking and how one handles it.
- The one traveling is a person, so information about her and her situation are useful. An example of this is knowledge about her ability to ride bicycles, which impacts recommendations given to her. In addition to information about the person herself, information about other people might also be useful. An example of this is information about the traffic on a road.

Traveling is, in simple terms, traversing a path. To travel on a road, information about it is important. Public roads in Norway are cleared of snow regularly and maintained by the government. Information about snow height and potholes can increase the efficiency of this and guide people away from roads that are difficult to traverse.

Traveling conditions are closely linked to weather conditions. Wind, sunshine, and temperature are only some of the weather conditions that can affect the transport method decision. A thicker jacket or an umbrella might be enough to change a decision to walking instead of driving. The measured temperature might be different from the apparent, or *human felt*, temperature [10]. Accurate information on how cold a person would feel outside can be used to give them better recommendations and improve their travel experiences.

All roads must come to an end and have a beginning. As such, information about the condition of parking lots can be interesting when looking into possible travel paths. An example of this is, knowing where vacant parking lots are. This can be used to reduce the amount of idle driving one has to do before finding a place to park.

When one travels, especially in traffic, one must prepare to meet other vehicles on the road. Different traffic conditions must be handled in different ways and can affect the wanted travel method. A prominent example of this is rush traffic. As more people can use the same roads, general information about cars on different roads can also be used to reduce traffic congestion.

There are some things that one must expect can happen on roads, but that can be hard to prepare for or predict. Extreme weather conditions and traffic accidents are examples of this. Getting updates about accident information might help in deciding what roads to take. An example of this is that extreme wind can close some bridges. Predicting this can help people plan so that they are not trapped on one side of a bridge. Statistics, information, and analyses can help with decision making in these situations if enough useful data is collected.

The information types listed above can be collected from different sources. A part of the goals of this project has been to investigate what sources to collect from and what can be collected. Data sources investigated for this research project can be categorized into **three main types**: *first-party*, *third-party* and *crowdsourced* data sources.

First-party data sources are sources deployed by or collected from directly by the UiT — The Arctic University of Norway. Temperature data collected from the weather station deployed by the university is an example of this. Third-party data is collected from organizations or businesses. An example of data collected from a third-party data source is downpour data collected from the Meteorological Institute (MET) in Tromsø[5]. Crowdsourced information is collected directly from people. User profiles filled in by a user is an example of this.

Some information can be collected from multiple sources and source types. An example of this is that the temperature in Tromsø can be collected both from the weather station at UiT and from MET. The trustworthiness and accuracy of the data must be taken into consideration when analysing data.

First-party data can be considered among the most trustworthy, as the information comes from ourselves. A problem with this information is that it can be inaccurate. This is because some of the data collectors used at the university could have been designed by students instead of professionals. MET weather stations are more accurate than student created devices, as they are designed, maintained and calibrated by professionals. Third-party sources have lower trustworthiness than self-provided data since we must trust them to provide us with accurate and valid data.

5. https://www.met.no/en

1.4 Methodology

This project has progressed as a type of prototyping and development procedure. The goal of the project was to provide information to users to reduce uncertainty in traveling. Little knowledge about the available information sources and their relevancy was known at the start of the project. An initial platform for future work was created as the project progressed. The methodology of this project can be described through some of the work of Phillip Glen Armour.

In the paper 'The Five Orders of Ignorance' [11], Armour explains about a different view of software development. The view is explained as regarding software as a medium of knowledge, instead of as a product. By shifting to this viewpoint, the focus of a software development project becomes acquiring knowledge, instead of creating a product. Armour talks about how software development has **five orders of ignorance**. The software can only be created when you have reached the lowest level. The orders start at 0 and stop at the 4th order. The orders of ignorance are, in order of lowest to highest: *Lack of Ignorance*, *Lack of Knowledge*, *Lack of Awareness*, *Lack of Process*, and *Lack of Meta Ignorance*. The orders start at 0 and stop at the 4th order

0. *Lack of Ignorance*: The designer has the required knowledge needed to solve a problem and knows how to implement a system capable of solving it.

1. *Lack of Knowledge*: The designer does not have the knowledge to solve the problem, but he knows what the problem is.

2. *Lack of Awareness*: The designer does not know what the questions to ask are. While he does not know what to ask, he knows that he lacks this awareness. This is often were a project starts out. At the beginning of this project, we had the goal of providing information to users that could impact their travel choices. What information to collect, how to analyse it, and how to present it eluded us. Before we could start answering question, we had to find what questions to ask.

3. *Lack of Process*: At this order of ignorance the designer does not even now what he does not know. In addition, time constraints make it impossible to acquire the awareness he needs to ask all the questions that needs to be asked. It is hard to determine what knowledge is in this order, as we lack an awareness of our lack of awareness. One way of solving this problem is to ask consultants what needs to be investigated.

4. *Lack of Meta Ignorance*: The designer knows of *The Five Orders of Ignorance*. We can use *Orders of Ignorance* to classify what we know and what we

don't know. This can be used to estimate the likelihood of what we don't know we don't know.

This project started out with some information in the 1st order of ignorance, some in the 2nd and some in the 3rd. The lack of knowledge about this system also placed us into the 4th order of ignorance, but we could still operate within the other orders. Placing constraints on the project moved us into a place where a higher percentage of knowledge was of 2nd order instead of the 3rd. This means that, while there was a lack of awareness about the knowledge required, it was not out of reach to acquire this awareness. The information gathered about knowledge in the 2nd order created questions in the 1st order that could be answered. Answering questions created knowledge in the 0th order. 0th order knowledge created software, a deployed IoT device and the foundation of a platform that could be the answer to this project.

1st and 0th order information acquired will be talked about in the Design, Implementation and Evaluation sections. The Related Work section will explain some of the information acquired to move knowledge from the 3rd and 2nd order of ignorance. We will investigate what 0th and 1st order knowledge was created by this project in the Discussion section. The new knowledge reduces the amount of knowledge in the 2nd and 3rd order. This will allow the project to progress further in the future. As this methodology looks at software development as a knowledge-acquisition method, the knowledge from this project will be able to further the work of future projects. Failures during this project can also be used, as knowledge about what not to do is also valuable information.

1.5 Thesis statement

The focus of this project is to collect, analyse and present information to lessen uncertainty in transportation brought by public and non-vehicular travel options. This is done in the belief that certainty in travel will increase the usage of environmentally friendly transportation methods.

A time limit of half a year restricts this phase of the project. As such, we limit the scope of the initial goals to be done. In this part of the project, we focus on collecting information from at least one of each of the data source types described earlier in this section. In other words, we collect some information from first-party, third-party, and crowdsourced data sources. The data collected will be analysed, either together with other sources or by itself, and will be presented to users. By collecting, analysing, and presenting the information we will have created a foundation for future work for this project.

In addition, we will investigate what other types of data can be collected and come with suggestions on how to proceed with this project after this thesis is complete.

1.6 Highlights

The following points highlights the main information found in this thesis.

- There is a lot of open data relevant for traveling that can be collected from different sources.
- There are multiple ways to detect traffic, weather effects, and other relevant data. The main issue is choosing what tools to use based on needs or resources.
- Weather seems to impact the distance sensor of the created IoT device, making information gathering error-prone during, and after, rainy periods.
- General advice and simple analyses can be done with small amounts of data and with general knowledge of human preference.

1.7 Summary

This project is about collecting information from multiple sources, analysing it, and then presenting it to users. The presented data should provide useful information about how it is or will be outdoors for the next few hours. By presenting information, we hope that it will reduce uncertainty in outdoors traveling. When users have confidence in their mode of transportation, they might choose a more environmentally friendly transportation method than their default option.

- Chapter 2 presents information about research into nudges and ways to collect data relevant for this project is described.
- Chapter 3 presents the general design principles underlying the project.
- Chapter 4 presents the actualization of the designs and the experiments done.

- Chapter 5 presents an evaluation of the experiments done.
- Chapter 6 presents an evaluation and discussion of the project as a whole.
- Chapter 7 concludes the project with a short summary and a look into work to be done in the future.

Now, we will move on to some of the research done for this project. Both nudges and information systems will be investigated in the Related Work section.

2

Related Work

A large part of this project is to collect data relevant to the observation of road conditions from multiple data sources. As such, there is a lot of different work that can be referred to as somewhat related. Some of these works will be written about in this section. In the introduction, four types of relevant data sources for this project were described. We will present some work related to these, together with work on nudging.

First, we will investigate what a nudge is, then monitoring methods for weather, roads, traffic, parking, and people.

2.1 Nudging and IoT

The Open Distributed Systems (ODS) group at UiT — The Arctic University of Norway is currently working on the NUDGE project, meant to help people choose green transportation options [12]. They do this by utilizing a tool called "nudging". This tool tries to change a person's behaviour without removing their freedom of choice. While the term was originally used in economics, other sciences and groups have started to utilize it. In Thaler and Sunstein [13], nudges are defined as:

"*... any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their eco-*

nomic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid." [13, p. 6]

Smart nudging is the next step in nudging [14]. From the definition above we can see that nudges have to preserve the choice of the one being nudged. Banning the usage of cars in the city is not a nudge, while giving motivational messages about using the bicycle in nice weather is. The step from a nudge to a smart nudge is being able to understand a person' needs. Research into the opinion on nudges shows that people have a mostly good outlook on nudges across nations [15, 16]. It also shows that different people are affected differently by the same nudge, which is also seen in other research [17]. Sociodemographic factors like age and gender affect people's choices, as do their culture. Smart nudging takes the fact that people and situations can be different into consideration. A smart nudge understands the user and his needs. While a nudge could be to give motivational messages about using bicycles, a smart nudge might do so when the weather is nice, and the user might have an appointment he can reach on the bike.

A nudge is comprised of **four tools** [18]:

1. *Simplification and Framing of Information*: Influence users by how information is displayed, positively or negatively, and using simple, easy to understand terms. An example of this is informing people about the positive effects walking has compared to driving.

2. *Changes to the Physical Environment*: Influence users by changing the physical environment to make the preferred choice easier to take. A nudge, in this case, could be moving fruits into more noticeable places while placing unhealthy food further away.

3. *Changes to the Default Policy*: Influence users' choice by changing the default choice to the preferred option. Changing a person's default policy travel plan from *ransportation by car* to *transportation by bike* can nudge a user who does not have a transportation preference.

4. *The use of Social Norms*: Influence users by showing what other people do. If the person performs worse than other people, they might change the way they behave to increase their performance.

In digital nudging [19], which is the way the ODS group is currently researching nudges, the first, third and fourth tools of nudging can be used. This way of nudging guides people's behaviour in the way information is presented. It is important to select the right information to display at the right time. An example of digital nudging can be found in [17], where digital nudging was

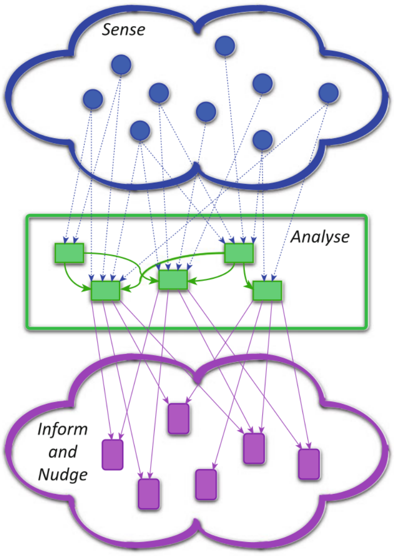


Figure 2.1: The publish-subscribe architecture with the *Sense*, *Analyse* and *Inform and Nudge components* — Taken from Andersen et al. [12].

used to help reduce the amount of incompatible online purchases made during online shopping. The ODS group combines the use of personalization, digital nudging, and situational awareness to make people choose green transportation options [14]. Changing the environment for digital nudging is possible, but not the focus of the group. An example of this would be to install screens on bus stops to display motivational messages.

We can see what affects a specific user and her needs by using personalization. An example of personalized nudging is removing nudges that cannot affect a user. To give an example, we can remove nudges towards the usage of bicycles if a user never learned how to use one. Situational awareness tells us what situation the user is in. A user that drives to work is not likely to be affected by a nudge that tells her to walk home. Both personalization and situational awareness can have use of IoT devices.

The smart nudging architecture that the ODS group presents uses three phases: *sense*, *analyse*, *inform and nudge*. Data can be collected from IoT devices, third-party sources, and users. This data can be analysed to get information about situational awareness and personalization. Analysed data can either be used in further analyses or to inform and nudge users. As an example, information from a phone can be collected and analysed to show how active a person is. A nudge encouraging more activities can be used if he is found to be less active than usual.

In Bothos et al. [20] a digital nudging project was tested out in Vienna and Dublin. Information about users and the environment was collected and used to promote green transportation. Eight different information types, together with a user's profile, was used to suggest using a car, public transportation or walking. While the research did not detect any changes caused by the persuasive messages, the overall response to the nudging was either positive or indifferent. The users expressed a desire for changing the presentation design. This could mean that lack of nudge impact was caused by wrongful usage of the first nudge tool, *simplification and framing*.

There has been some concern towards the ethicality of nudging [15,16,18]. One of the reasons for this is that nudging has been a tool that tries to affect the subconscious of the users. While users could be told about the nudges it has been believed that this could negatively affect the users, sometimes making them choose another option out of spit. In Bruns et al. [21] transparent nudges were tested with regards to environmental friendliness. In the research, it was found that telling people about the usage of a default nudge did not greatly impact their choices. The usage of transparent nudges was also not seen as freedom threatening or a source of anger.

2.2 Weather

The weather now, yesterday, and the forecast for the day can all affect the travel decisions of people living in an area. In this subsection, we will investigate work on cataloguing weather statistics, weather forecasts, and collection of weather data.

Firstly, three providers of meteorological data will be presented. Afterward, some different ways of measuring the effect of the weather on the environment will be looked at. The providers of meteorological data were selected based on what information can be collected about Norway, as this thesis is being done there. Two of the providers are institutional, while the last one is a firm. The institutional providers are the *Meteorological Institute* in Norway and *NASA*, while NETATMO is the firm we will look at.

The Norwegian Meteorological Institute [22], or MET, was established in 1866. Today it is one of the leading international centres of expertise in meteorological studies. All their work has as its main goal to help protect life and property. From their many APIs, both historical and forecast data can be collected.

2.2.1 Weather data Providers

NASA, the *National Aeronautics and Space Administration*, is an independent agency of the United States Federal Government [23]. A part of their research goes into collecting weather data and it is possible to look at their collected values on their websites [24].

NETATMO is a firm where one can buy different home monitoring solutions [25]. One of their product lines is weather stations. People can buy a weather station and collect data like temperature, humidity, and air quality. This data can then be uploaded to a common database. One can look at the data from weather stations at their weather map service or extract the data through an API [26].

2.2.2 Effects of Weather

The effect of weather can be measured in different ways. Examples of this would be using a humidity sensor to see if it has rained. Another example would be to use a snow sensor to measure snow height and snow conditions. This was done in Svalbard for Longyearbyen [27].

Mud is one of the effects of the rain that can have an impact on the decision of traveling by non-vehicular means. When there is mud on the road some bikes have problems traveling and summer shoes might have water seepage. There are several ways to detect mud, but it can be hard to differentiate mud from other road anomalies. Mud is one of the things that are easiest detected by their effect on their surroundings instead of directly measuring it. This is because it can be difficult to measure the actual water content of soil [28].

A problem with accurate measurement of Soil Moisture Content (SMC) is that one must calibrate the sensor analysis with the specific properties of the earth. As earth is a collection of different materials, different compositions can alter the readings done upon it. During a field evaluation and performance comparison of soil moisture sensors, different sensors were tested in an agriculture setting over several years. Uncalibrated sensors were accurate enough to create sensor specific low/high levels of water content to show when to start or stop irrigation [29].

In addition to ways of collecting information directly from the soil, like measuring the electrical resistance of the soil, one can detect SMC remotely in different ways. In 1995 research was done on using microwaves in remote sensing and great initial results were found [30]. Several experiments using sensors mounted on trucks, aircraft, and spaceborne sensors showed that a thin layer of the soil could be accurately measured. This means that finding mud, which is on the surface layer of the ground, should be possible through airborne drones and other units sensing remotely. Here again, detecting changes or differences in the soil is easier than accurate values.

JPL Robotics has also done some research on remote detection of mud. They have focused on mud detection during optimal conditions. The research was done to find sensors that could help guide unmanned ground vehicles through rough terrain, as mud could stop Unmanned Ground Vehicles (UGV) in their assignments. In worst case scenarios the UGVs could be lost or stopped in a place where they could not be extracted. This means that the UGV is effectively lost [31,32]. Examples of mud detection conditions are that mud only occurs on the ground surface, is cooler than dry soil during the daytime under nominal weather conditions, and it is highly polarized. JPL collected data on wet and dry soil in many ways and characterized their advantages and disadvantages.

These options tested by JPL include colour cameras, NIR vs Red reflectance and polarization of the light reflected by mud. The separation of mud from other terrain obstacles was also researched. NDVI can separate soil from vegetation, while DASL can separate wet and dry soil. From their research, they implemented a mud detector with multiple imagery sensors. Using this detector, they managed to spot mud and map it to a created world map.

2.3 Road monitoring

A lot of research and work has gone into determining the conditions of roads. Specialized vehicles, vehicle add-ons, and static solutions have been tried and some will be described in this subsection. Quick and cheap detection of road anomalies can save lives in traffic and decrease the costs of road maintenance.

The *Observator* group develops and delivers sensors and systems for meteorological and hydrological monitoring. Some of the information related to road monitoring allows for the measurement of variables like precipitation levels and snow depth. Examples of different road conditions are cracked roads or snow-coverage. Their system provides everything from telemetry to data processing.

Roadstar [33] is another finished solution in the form of a specialized monitoring vehicle. It comes with a wide range of sensors and capabilities. At 60 km/h it can measure entire lanes for cracks. Overall, the vehicle can look for a wide arrange of road issues, like cracks or ice. Together with a large amount of actual road monitoring, information about parts of the road environment is also collected. Examples are the positioning of road markers and 3D stereo imaging of the lane width.

Some work has been done to equip vehicles with additional sensors to provide road monitoring. Examples of such systems are *RoadMonitor* [34], *Pothole Patrol* [35] and *BusNet* [36]. *Pothole Patrol* and *BusNet* equipped taxis or buses with accelerometers to detect road anomalies. *RoadMonitor* furthered this research by using the gyroscope and accelerometer of smartphones to the increase detection accuracy. Road anomalies can be detected or inferred from when vehicles lower their speed or when there is a sudden spike in upwards movement. While there are some drawbacks from this type of detection, like false positives, the usage of this type of detection decreases the need for specialized vehicles. Equipping taxis or buses with sensor equipment allows for monitoring of many roads with a small number of sensors. This comes from the fact that taxis and buses are mobile units that can monitor their areas or routes continuously.

Another system for participatory, crowdsource sensing is a work from 2012 [37] that created an application for using the in-vehicle smartphone of vehicles. Modern vehicles have a lot of sensors inside them to monitor their condition. The finalized application was going to be able to classify potholes, rough roads, and bumps. By accessing the vehicles GPS, accelerometer, magnetometer, compass and connectivity modules, the application could collect data, analyse it and transmit the analysis to a backend that aggregated the reports. The application was client heavy but was both secure and protected user privacy. In addition, it was both cost and energy efficient.

Vehicle-to-vehicle communication can be used in road monitoring when normal communication is difficult. The *SODiCS*[1] system investigates this [38]. It is meant to be able to propagate useful information in disaster areas to help in evacuation and rescue situations. Data propagation becomes possible by turning vehicles into mobile, Internet of Thing units. Each vehicle can collect, transport and transmit data wirelessly to and from data stations and other vehicles. Vehicles monitor the road and transmit useful information to passing cars.

One road monitoring system used cars with connected sensors as mobile sensing units and used the driver's phone to transmit the collected data [39]. Like *SODiCS*, the vehicles transmitted data to and from base stations. The sensing device used an accelerometer, like in *Pothole Patrol* and *BusNet*, to locate road irregularities. Together with the accelerometer readings, audio

1. Spatial and temporal Omnidirectional sensor data Distribution and Collection System

data was collected to analyse the pavement conditions. The analysed road condition was given a score about how serviceable the road was, which was then used to rate roads.

The last system to be described under road condition monitoring systems is one created for the Indian road network [40]. Many assumptions for a developed country's traffic system was not applicable when the article was written. Limitations were present because of the poor traffic system at the time and the low amount of resources available. As the main contributions, the system was going to be able to offer alternative routes to cars on congested roads and areas of frequent congestion. In addition, the system should spot potholes to quickly and cheaply do road maintenance. To do traffic analysis a metric called "traffic density" was created. It used magnetic readings to find the amount of traffic in a lane. The metric showed an aggregation of all the cars on a stretch of the road instead of a specific vehicle count. This was done to simplify the monitoring system. Accelerometers were once again used to detect potholes.

2.3.1 Traffic

Monitoring traffic can help when deciding whether to travel by non-vehicular means or with public transportation. Different ways of monitoring traffic can be split into two groups: static and dynamic. The criteria for if something is static is that the detection unit is focused on a specific spot or area. Dynamic monitoring units move with the traffic. Some static and dynamic methods of traffic monitoring will be described in this subsection. Information about both static and dynamic methods of traffic monitoring has been collected by the Joint Research Centre [41].

In the previously described road monitoring network from India, they used a static approach to measure a metric they called traffic density. The system used multiple magnetic sensors to be able to scan for how much of the road was occupied by vehicles. [40]. Another way to statically detect traffic can be done by using cameras. Images of an empty road and the same road with cars on can be compared by computers. By analysing the differences, it can find out if there is something on the road and how many separate somethings there are. There is already open-source code for this, which means anyone with a computer and a camera can do it [42]. Other options include using laser tripwires or sonars to detect cars on a specific stretch of road.

Dynamic ways of collecting traffic data generally fall under *Floating Car Data*, or FCD. The principle of FCD is to collect and track real-time traffic data through GPS data or mobile phones inside a vehicle. Any vehicle with either, or both, can then be turned into a mobile sensing unit of an intelligent transportation system. By taking in the location, speed, and direction of a vehicle, an estimation of the congestion can be made. Multiple vehicles transmitting such data can be used to find the vehicle density of an area. This can be used to locate congestion. Data from vehicles can be collected and transmitted anonymously to protect privacy. Aggregated data can then be processed on servers. The processed data can also be transmitted directly to cars that connect to the server. Getting FCD from mobiles does not require any additional hardware to be created. An inconvenience of using mobile phones to get FCD is that the accuracy is low, with a typical precision being around 300-meter. This weakness is compensated by having many devices to collect data from.

Floating Car Data is already in use by multiple organizations and applications. Google maps [43] is one such application. The application is used to help in travel planning between two spots. One of its features provide information about the amount of congestion on roads.

Tracking of buses can also be used as FCD and can be used to get information about the traffic flow of specific roads at given times. A travel helper for Troms county in Norway provides updates about how delayed a bus is [44]. The lateness number increases as the bus gets further behind schedule. This information is also possible to extract through API calls. Traffic flow can be inferred by looking at how this delay changes the time increases. Troms county is also working on ways to allow the actual location of buses to be extracted, which can also help in monitoring traffic flow through publicly available sources.

A part of helping people choose greener traffic options is to lower the bar of finding easy travel routes. Applications like the Troms county bus travel planner is an example of such a solution, as is the Google maps travel helper. Another example is "Entur" [45] which is a platform whose mission is to enable and increase the use of public transport in Norway. The application collects data about different public transport options and allows for the purchase of railroad tickets to get to different places. "Entur" collects data from 60 public transportation operators and has a registry that contains data about 21,000 daily departures on 3,000 routes. All this data is open and free to use for app and service developers.

2.4 Parking

Organizing and locating parking has been dealt with in several different ways. Most of the time, one simply must go to the parking lot to see how many lots are left. In this subsection, we mention shortly some of the ways to monitor parking spaces and how to make the parking process easier.

The company *Windmill Software* has compiled a list of different ways to detect cars [46]. Some can be used in both traffic and parking situations. Counting the number of cars manually, camera imagery and different types of tripwires are only some of the ways to detect cars. Cameras can be used to check for parking vacancy by comparing a picture of an empty lot with a current picture of the same lot. One can also use image recognition to find the number of lots in a parking space that are vacant.

Parking applications can be used to locate possible parking destinations [47]. After arriving at a parking area, the area can use an indication system to guide cars to empty lots. These indication systems can be connected to sensors that check if a lot is empty. An example of this is connecting detection sensors to overhead lights in underground parking spaces. When a light is on, the parking lot underneath is vacant. This makes it easier to find a spot quickly.

2.5 People

Information about, or from, people can be collected in multiple ways. An example of an intrusive way of collecting information is a survey that people must fill in, like a form describing health issues. A non-intrusive way of collecting information is a health app on a phone that counts the number of steps taken in a day. Some ways of collecting information about people will be looked at in this subsection.

Health applications come in different forms and collect different information. Some applications come with additional hardware options, like smart watches to measure heartbeats or blood pressure. This information can be used to tailor experiences for getting people to be more active, by allowing an inferred health degree affect recommendations on travel methods. An example would be that an active person might be less affected by weather, while a more inactive person might only like to walk in nice weather.

People can provide useful information about their health and about the roads in a descriptive fashion. Posting information on social media, like on *Twitter*, can allow the public to acquire such information. An example of this is people posting *tweets* to the police or to the local traffic monitoring authority.

Some specialized sites and applications are used to describe the roads allow people to post information about roads for either the government or other

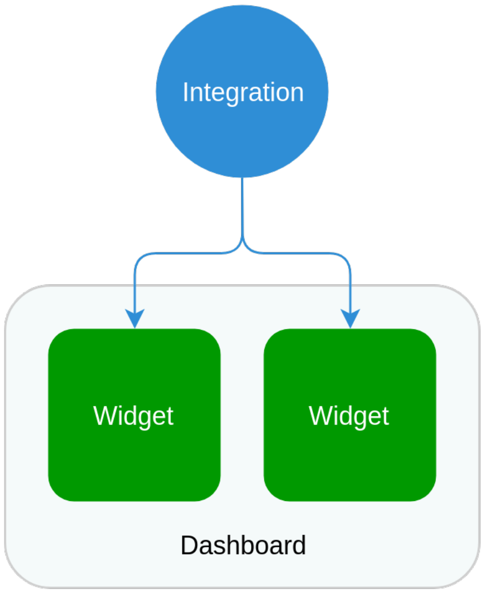


Figure 2.2: VisualBox Model [50]

people to see. Examples of this are *Gatami* [48] and *Google Waze* [1]. *Gatami* is a site where Norwegians can, among other things, post issues about their roads. An example of such an issue could be that the road has been damaged. *Google Waze* allows drivers to post useful, driving-related information to other drivers.

2.6 Visualization

VisualBox was and is being created for the purpose of testing visualization options in a relatively easy way before developing long term solutions [49]. The application is an integration and visualization tool that runs mainly in your web browser. As the applications run in a web browser the only thing needed to use it is an internet connection and an account.

The application consists of three main parts: *integrations*, *widgets*, and *dashboards*. An integration is used to fetch information, process it and to generate data models. Widgets are the visualizations tools that present the data to the users. The dashboard is where the widgets are placed to visualize the data.

When an integration is created one can decide what type of programming language it is supposed to use. The program then sets up a Docker container running with the runtime chosen and runs the code provided. Currently, at the time of writing, the runtimes supported are: *Golang*, *Node.js*, and *Python 3.6*. Code for both widgets and integrations can be imported and exported to allow for local storage of the code. Widget creation currently only support *JavaScript* runtimes.

Integrations and widgets can be published to VisualBox to allow others to use the same code. Some currently, publicly available integrations allow people to track satellites using *Space-Track* [51] data or collect weather data from the weather station at UiT.

2.7 Summary

This section talks about nudges and ways people have, or can, collect and present information that could be useful for this project. Nudges try to non-forcefully change an affected party's behaviour. The ethics of nudges and some tests done on them are mentioned in this section.

A system nudging people towards environmentally friendly transportation was described, which will, in the Discussion section, be compared to the system created during this project. Of note is that the system created for this project will aim to collect a lot more data than the system described in this section. In addition to collecting more information from outside sources, and of a wider variety, this project will also supplement with data collected from self-hosted sensors and devices.

In the following section, Design, we will investigate the initial designs for what was going to be done in this project.

3 Design

The design of this project is divided into two main parts, information collection and analysis, and information presentation. Information is collected from UiT — The Arctic University of Norway, the Meteorological Institute (MET) and from people. The information is processed and presented onto *VisualBox*, a visualization software created by a student at UiT [49]. Data is collected from weather stations hosted by both UiT and MET. Information about the roads is collected from an IoT device created and deployed for this project as information about snow height. Notifications about road issues are collected from users over *Twitter*.

3.1 General Architecture

A *data warehouse* approach is used as the general design for this project. There are three components in a data warehouse approach: *sources*, the *processing unit*, and *endpoints*. The processing unit of a data warehouse has three main steps: *Extract*, *Transform*, and *Load* (ETL). Information is extracted from UiT, MET and from users. The information extracted is transformed into similar formats and analysed. After transformation, the information is loaded into a specified endpoint. In the current design, both the ETL component and the final endpoint is the *VisualBox* tool. The data is extracted, transformed and loaded by VisualBox *integrations*. VisualBox *widgets* are the final endpoints, where the data is presented to users.

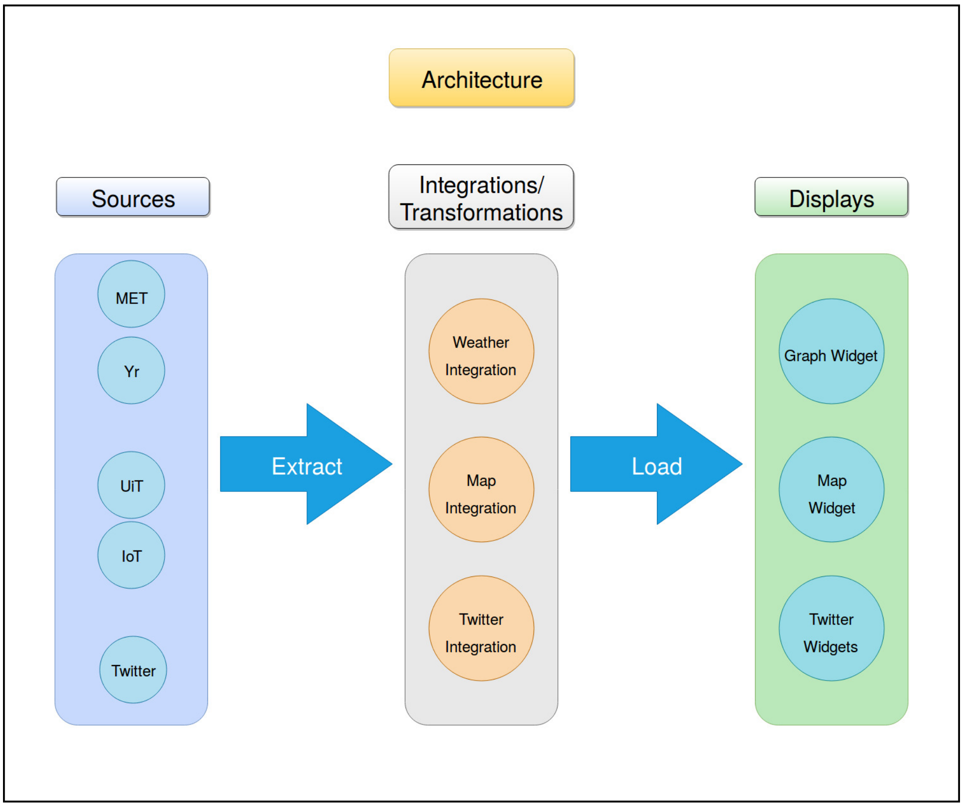


Figure 3.1: The current general architecture of the *RoadAhead* project.

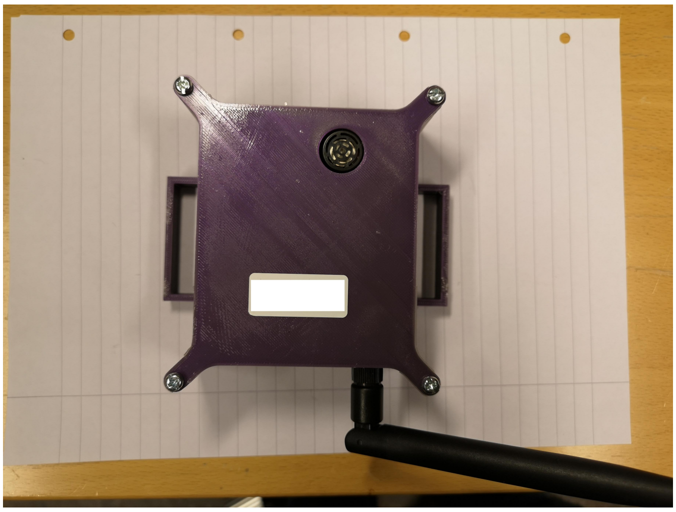


Figure 3.2: The current iteration of the IoT device meant to measure snow height.

3.2 Collection

The designs of how data was collected will be presented in this subsection. We divide this subsection into how we collect similar types of data. Data about snow heights and some auxiliary information was collected from a deployed sensor. Weather data was collected both from the UiT and from MET. Crowdsourced data was collected over Twitter.

3.2.1 Deployed sensor

The design of the deployed sensor can be divided into three parts: *hardware*, *software*, and *casing*. An initial design was taken from previous work done on a snow sensor [12]. That device also used a sonar, distance sensor, but was designed to work on while connected to a large car battery.

The deployed sensor is going to measure snow height as its primary job. In addition, it will collect supplementary data like temperature, humidity, and acceleration. Snow height can give snow plowers a more accurate way of knowing when they must plow. Temperature data and humidity can allow people to know what type of snow they can expect on the roads. Dry, cold snow might cause powder snow, while high humidity and warm temperature can cause tightly packed snow. Acceleration data can allow maintenance personnel to know if the device has been disturbed by the environment. An example this is that wind can tilt the device, which will cause changes in acceleration for the device. A tilted device would provide false data readings to the backend.

The deployed IoT device uses a microcontroller, a distance sensor, an antenna, a specialized sensor extension board, and a battery as its hardware components. Information is collected through the sensor by the microcontroller which runs the device. This data is then transmitted over the LoRaWAN [52] network to a backend server. The usage of an antenna should allow the device to send data over a long range. The extension board has additional sensors that collect auxiliary information about the environment around the device. A battery is connected to the microcontroller through the extension board.

To measure snow height, a distance sensor is used. The initial distances collected from a device can be used as a reference point by the device. Any increase or decrease in subsequent measurements will mean that the height of the snow has changed. If the distance to the ground is known from the beginning of, the difference between the measured height and the real height can be used to determine the snow height.

The finished version of the IoT software can be divided into three main parts: *sensing*, *sending* and *sleeping*. Data from the distance sensor and the extension board is collected in the sensing part of the program. Multiple readings from the sensors are collected and the data is filtered to remove unrealistic sensor data. The leftover readings are averaged, being both temporarily stored and transmitted to the backend server. A duty cycling procedure is set in place to conserve power. How long each duty cycle lasts depends on the last few averaged distance readings. When larger differences are found, the duty cycle is shorter.

The casing of the IoT device is a 3D printed cuboid with a lid fastened with screws. Handles were designed to be on the sides of the box to allow the device to be fastened outside. As the components are stacked, with the battery at the bottom and the sensors and microcontroller on top, the casing is both tall and wide. The finished casing had the circuit board resting atop a plateau inside the box. A hole was made in the lid so that the distance sensor could sense outside the box. The antenna is designed to stick outside the box from the side. As the IoT device had to survive in bad weather conditions a protective spray was used to coat the casing.

3.2.2 Testing

The created IoT device was tested while in various environments and ways of deployment. Tests were performed both inside and outside. While outside, the device was also tested in various weather conditions. Information of interest was: the *uptime* of the device, the *accuracy* of the readings, if *bad weather was a limiting factor*, the *impact of water content* in the snow. Depending on the

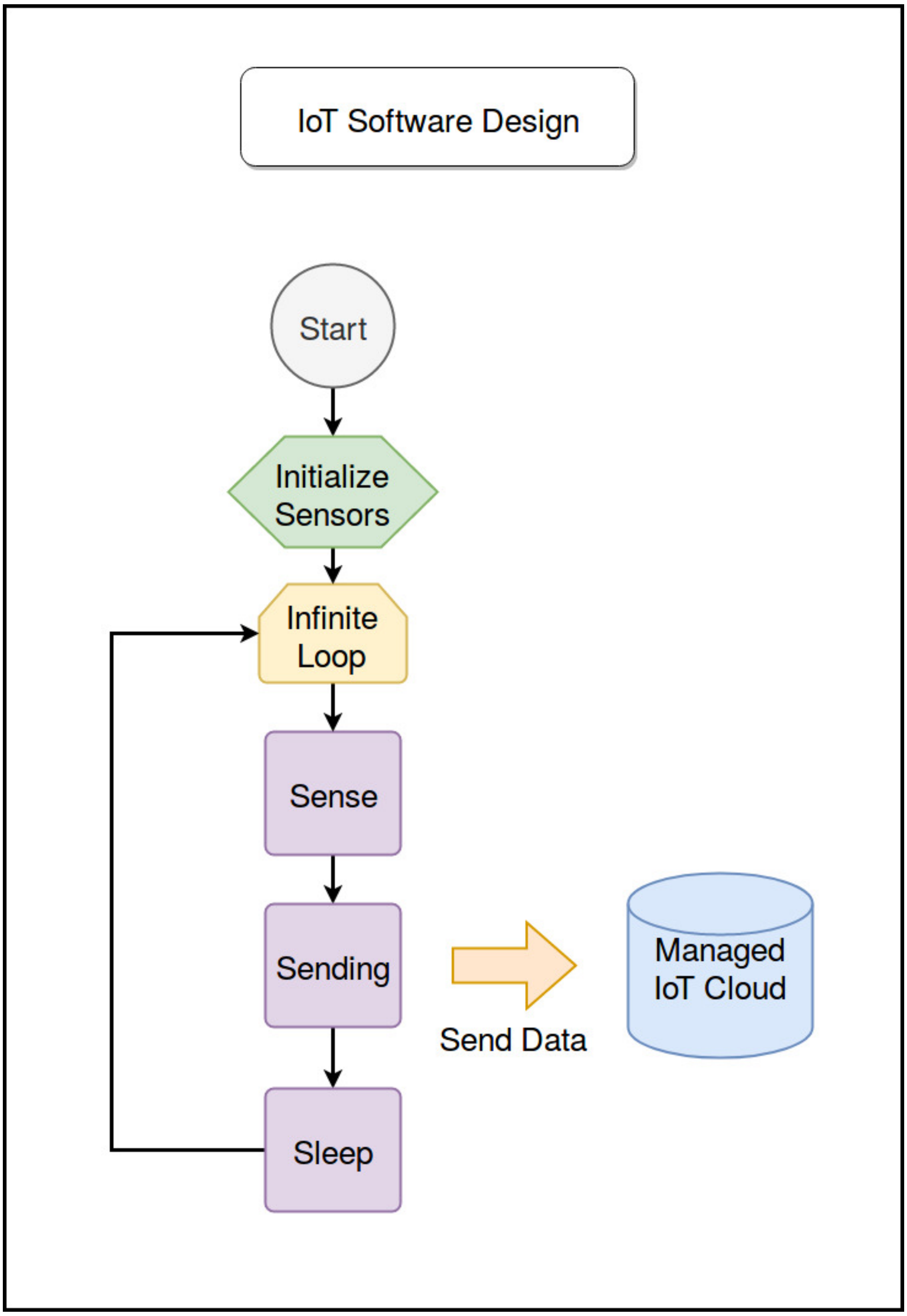


Figure 3.3: The general concept for the software of the IoT device.

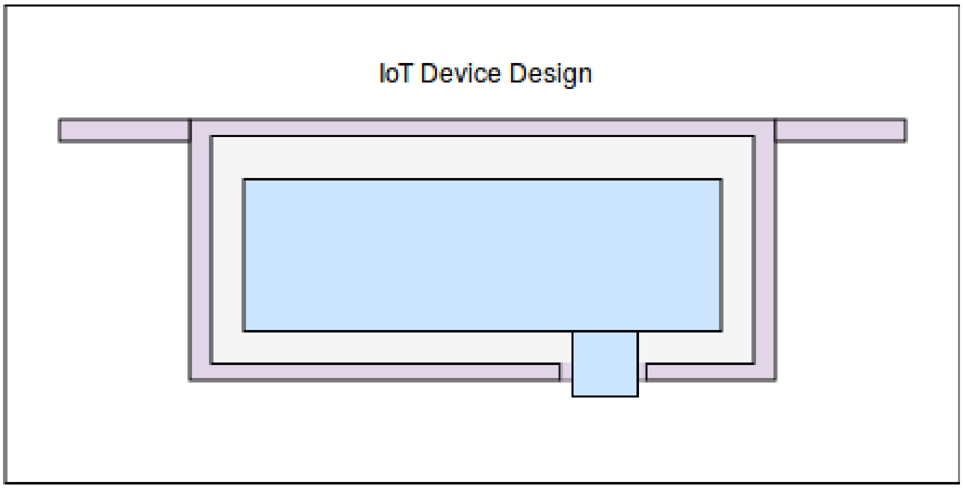


Figure 3.4: Concept of the casing with an opening for the distance sensor at the bottom. The antenna was not included.

results gotten from the tests, the viability of the device as a sensible way to measure snow height will be discovered. Some factors can have an impact on the device design, like having to exchange the battery to increase the uptime. Other factors can make the device unusable for the intended purpose, an example being that bad weather stops the device from being able to measure in a satisfactory manner.

3.2.3 Weather data

Three main sources of weather data are collected, two historical datasets, and one forecast. MET has many different open APIs with various data about the weather in Norway [53]. Included in these APIs are ways to get information about air quality, lightning, and tides. A part of the data collected by the deployed IoT device can be used as weather data, as the device measures the temperature and humidity.

The MET *Frost API* is used to collect historical weather data from the institution's weather stations [54]. The Frost API provides free access to an archive of historical weather and climate data. Some of the information includes hourly data logs from various weather stations. To make it easier for programmers to extract relevant data, the API allows users to find weather stations near a given location. In this way, one can locate and extract data from sources near a wanted location. At the time of writing, the Frost API is still being updated, with the latest update being from 14th. of April 2019.

The other historical dataset used is collected from a weather station at UiT. Historical weather data in the database dates back to 1993. This data is being collected at a per minute basis and includes information about wind strength, temperature, and downpour [55].

Forecast data is collected from *Yr*, which provides forecast data from MET in an easy to understand manner. Yr provides forecast information from all over the world. As part of their forecast, one can find an estimate of the temperature, wind speed, and downpour at different times of the day. Additionally, one can get different types of forecasts. Examples of these are hourly forecasts and daily forecasts. In this way, we can select the forecast wanted based on how long into the future we want to go. The hourly forecast is more accurate than the daily one but does not go as far into the future [56].

The main information collected in this phase of the project is information about downpour and temperature. The reason for choosing these data types is that weather can have a large impact on people willingness for traveling outdoors. An example of this is, that a little amount of downpour might allow the usage of a bus but negates the choice of walking. Furthermore, a large amount of rain might be fine during warm periods but will be a deterrent in cold ones.

3.2.4 Crowdsourced Data

Users can provide relevant data either about themselves or about the environment around them. Data about users can be used to tailor advice, while data about the environment can enrich a database with user verified information. An example of tailored user advice could be to lower the warning thresholds of slippery roads for users in advanced age. Giving users a good travel experience is paramount to have him continue to use the application. As he continues his usage, he can be nudged towards environmentally friendly transportation options when they will provide a good experience.

Humans are equipped with a plethora of sensors and can enrich the data they collect with the experience they have. This allows users to collect, analyse and transmit knowledge they find useful to wanting recipients. A problem with collecting user data is that human-readable text is not something computers can easily understand. This problem has been investigated in other works, an example being the work on the semantic web [57]. Translation from human-readable text to computer readable information needs to be part of the design.

To simplify the semantic part of the collection, a predetermined schema for messages was created. A widget was created to help users with posting information as a valid post. This was done by providing an easy way of transforming provided information into a valid format. The widget creates a link that can be used to post the information under the user's account.

Twitter is used in the current part of the project to allow users to provide data about the environment. The platform has APIs available to allow users to post and collect *tweets*. A *tweet* is a short message of 140 characters or less, that can also contain media like pictures. Tweets can be tied to a topic or another account through something called a *mentions*. *Mentions* ties content to a specific user. This connection can be used as search parameters in the *Twitter Search API*. In the current project, a Twitter user was used to tie the messages posted together.

The collection part of the setup periodically collects the latest posts that have a specified *mention*. As the posts comes in formatted in a readable way, the program can parse the posts and distribute them to the rest of the setup.

3.3 Presentation

Presentation of collected work and analyses will be done through *VisualBox*. A standalone platform can be made in the future. Visualization on VisualBox is done through *dashboard* and *widget* components. As such, we will investigate them for the presentation designs.

A dashboard can be considered a tab on a page, while widgets are separate, individual windows that can display data and can be interacted with by users. VisualBox allows users to change some configurations on dashboards and widgets through side panels. While the dashboards have a static configuration board, widgets have user-created configuration options. This allows users to configure the same widgets to look differently or display different information based on user input.

Four main displays of information are to be created: a *common map*, a way of *giving advice about the road*, *weather information* collected from MET, and the *formation and presentation of Twitter* messages.

The common map will contain information collected from weather stations, Twitter, deployed sensors, and weather data collected from MET. One of the main functions of the map will be to show temperature changes in an area. Weather data collected from MET will work as a type of *Global* value, as MET will have the most accurate and general weather data readings. Data collected from the weather station at UiT and from the deployed IoT device will work

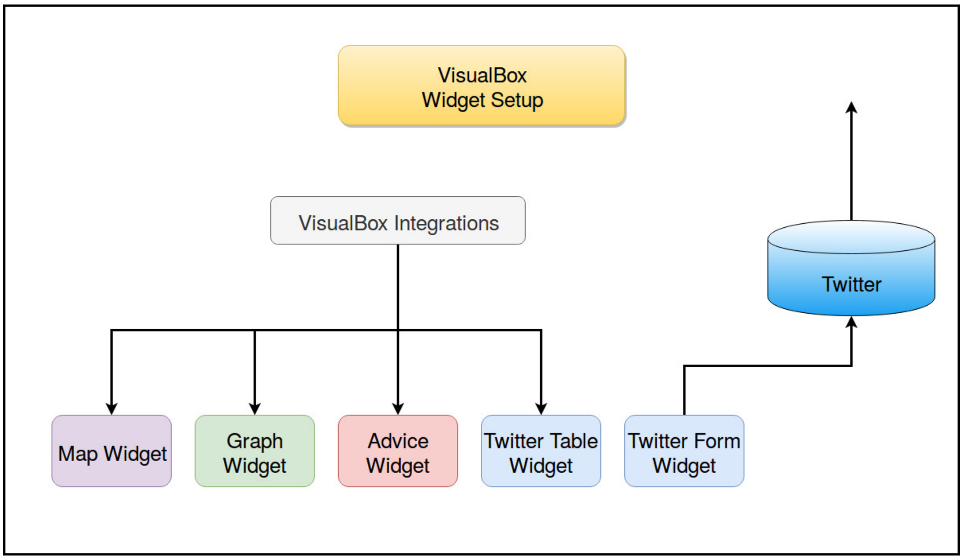


Figure 3.5: The planned setup for the widgets.

as *Local* values. The differences between a local value and the global value will be displayed on the map as coloured, overlapping circles. So, if the global temperature is 0 degrees and the local value is 10 degrees, the circle will represent 5 degrees. In the initial phase, only the temperature readings will be used, but multiple overlays can be made in the future. In addition to the temperature overlays, information extracted from Twitter will be added to the map as dots containing additional information when interacted with.

The weather widget should show the information about the last 12 hours and the future 12 hours of meteorological information. The most important information displayed should be the temperature. The temperature will tell people to dress in more clothing, which can take time to prepare. A case could be made for choosing to display the apparent temperature instead of the actual temperature. In this case, the temperature would be analysed together with the wind to give a better view of what one can expect to encounter during the day.

Advice should be given with several levels of abstraction to allow users to quickly and easily understand the message. This provides uncertainty of the number of widgets to create for this information. In addition, the information can be given about the present situation, or about the future. If a single widget is created, it will provide advice about the present and the future, and with several levels of abstraction. If several widgets are created, the information can be split into smaller pieces.

We will have two Twitter widgets to display information and interact with users.

The first widget will be a type of form that allows a user to enter information to post. As the widget takes in the information it can transform it into a computer-readable format. A second widget displaying the latest posts will likely be in the form of a table. As both forms and tables have a set number of fields, the method of posting and displaying information will complement each other.

3.4 Summary

In this section, we looked at the general designs of the components to be created for this project. Plans for what data to collect and how to present analysed data was described. Designs for an IoT device was looked at, together with its planned software procedure. In the next section, Implementation, we will investigate what was created in detail and how tests of the created IoT device was set up.

4

Implementation

In the following section, information about the actualized collection methods and design choices will be described in more detail. Firstly, the collection methods will be presented. Then, the presentation decisions of the collected data will be looked at.

4.1 Collection

Decisions about the different collection methods and integrations will be described in this subsection. The IoT device deployed in this project collects raw data and sends it to the *Telenor Managed IoT Cloud* [58], an IoT platform, and application toolbox provided by Telenor. The information is then transferred to VisualBox.

VisualBox, the visualization tool used in this project, collects data from outside sources, analyses them and presents the information to users. In this subsection, the different integrations implemented to analyse information will be described.

Integrations created were written in *NodeJS*, a programming platform that is used in web development. Most of the actual programming was done in the *JavaScript* programming language.

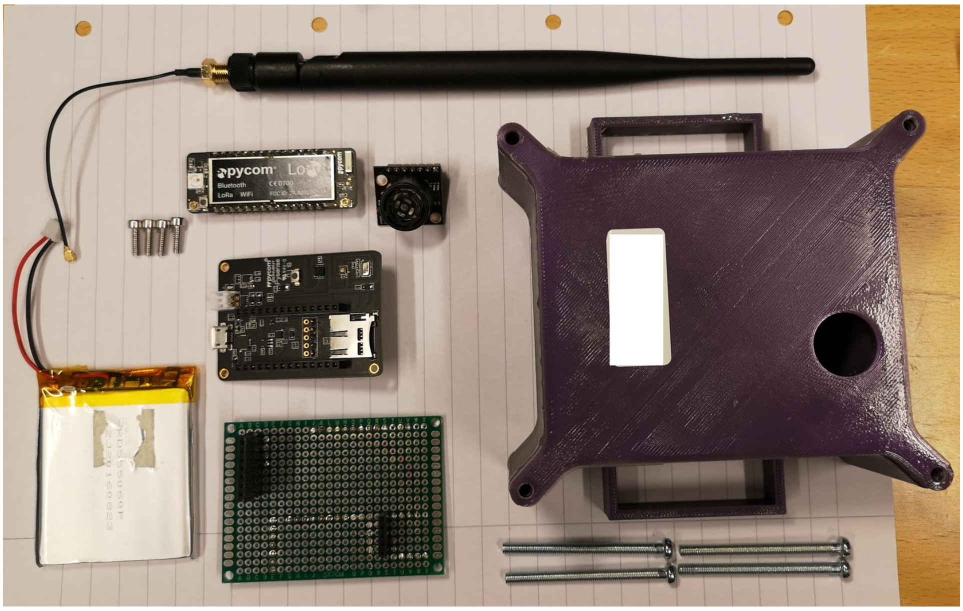


Figure 4.1: The parts used to create the IoT Device.

4.1.1 Deployed Sensor

In the following subsection, the deployed IoT device is described in further detail. After the description of the parts used and the casing created, the tests done will be listed and described. In the end, the integrations created to extract the data will be looked at.

Two Pycom components were used in the device, a LoPy microcontroller, and a Pysense extension board. Pycom provides several IoT solutions, together with *Do It Yourself* IoT parts. A LoPy microcontroller [59] was used to control the device because it can run the *micro-python* programming language and can transmit data over the LoRaWAN [52] network. The Pysense [60] extension board was used to allow a battery to easily power the LoPy device. In addition to the recharging capabilities, it comes with a sensor suite allowing acceleration, temperature and humidity data to be collected.

A *MaxSonar MB1043* distance sensor [61] was used to measure the distance to the snow surface. The sensor is supposed to have precision down to a few millimetres if a compatible temperature sensor is connected. Even if a temperature sensor is not connected directly to the sensor, it should still be accurate enough for the purpose of measuring snow height. A temperature sensor was not attached directly to the sensor during the testing. Its measuring range was between 300 to 5000 mm. As the device has a range of up to 5 m the IoT device could be fastened to trees or tripods above the snow.

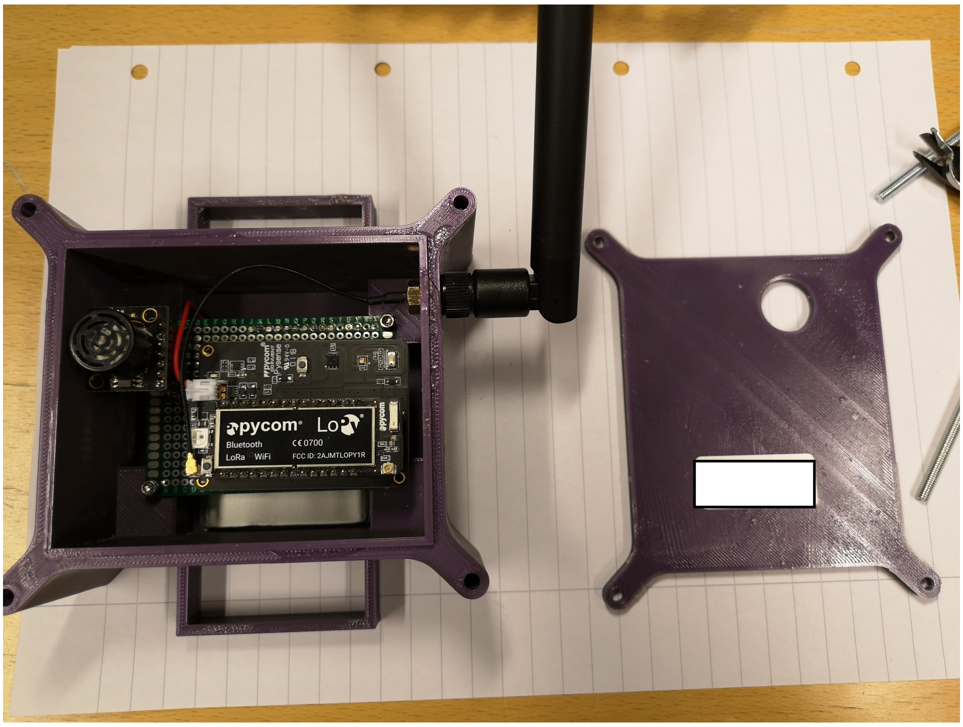


Figure 4.2: The finished IoT device. Opened to show the components.

A LoRaWAN antenna was connected to the device to increase the transmission range. The antenna should be enough to connect to the LoRaWAN network in the local area set up by Telenor. Their setup uses three base stations together with a few sporadically deployed smaller stations. All base stations are mounted on masts with relatively optimal placement. The first is placed in *Tønsvarden*, the second is placed on *Røstbakktoppen*, and the last is placed atop a mountain in *Tromsdalen* [62]. By using LoRaWAN to transmit data, small messages can be transmitted at low power cost.

The software used to run the IoT device has an initialization phase and the main program loop. In the beginning, the IoT device connects to the LoRaWAN network and initializes communication with the sensors. In the subsequent main loop, data is collected, processed and transmitted to a backend server.

Data from the distance sensor and the extension board is collected in the first part of the main loop. Multiple readings from the distance sensor are collected and large deviances are removed following the IQR statistical spread [63] method. The remaining readings are averaged and temporarily stored. After collecting data from the distance sensor, information from the extension board is collected. The expansion board is collected from multiple times and averaged, but no filtering is done. The processed information from the sensors is transmitted to the Managed IoT Cloud of Telenor.

A duty cycling procedure is set in place to conserve power. The main loop of the program switches between resting for 5 and 20 minutes. The time to rest depends on how large the difference of the last measured distances was. Therefore, distance results are temporarily stored. A shorter rest is used in periods of larger differences. Collection from the Pysense happens with a minimum interval of 30 minutes, as the main information to be collected comes from the distance sensor.

The device has two partly unnecessary power drains. Firstly, the device still uses power as if it was idle while the microcontroller is resting. This occurs because the LoPy does not have *deep sleep* capabilities, which would allow for low power resting. Secondly, the microcontroller does not deinitialize the communications with the sensors before sleeping. This takes a continuous, small amount of power. The device does not do this because of convenience.

The casing of the deployed sensor is a 3D printed cuboid with extended corners to allow for thinner walls while securing the lid screws. A handle on both long sides of the box was designed to help with fastening the device upside down during deployment. At the bottom of the casing, a depression was created for the battery.

A wide and tall casing was created for the device. The created casing was wide because of the use of a large battery and a wide circuit board. A large battery was chosen to allow for a longer deployment time, while a wide circuit board was chosen to allow all the components to be on the same board. As the components are stacked, with the battery at the bottom and the sensors and microcontroller on top, the casing is also tall.

The finished casing had the circuit board resting atop a plateau and fastened with screws. Under the circuit board was the battery. The Pysense and the distance sensor was fixed atop the circuit board. On top of the Pysense, was the microcontroller. A hole was made in the lid so that the distance sensor could sense outside the box. The antenna was designed to stick outside the box, through a second hole made on the side of the casing. A spray was applied to the IoT device to give it a layered protection. This protection allows it to survive in bad weather conditions. The layered protection was applied by lacquering the casing multiple times.

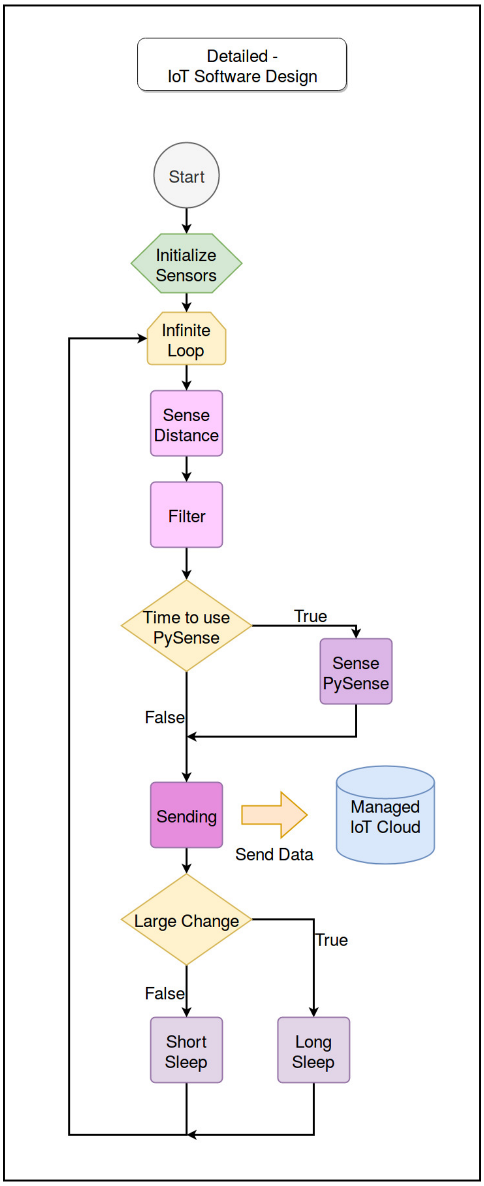


Figure 4.3: A more detailed view into the software dataflow of the deployed IoT device.

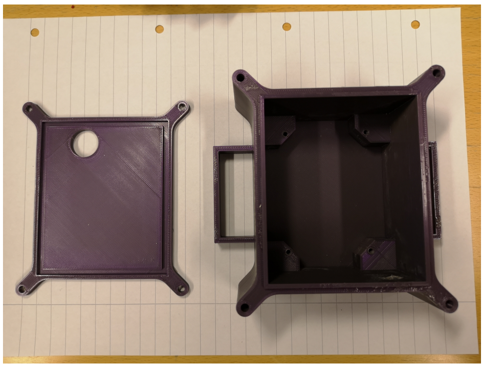


Figure 4.4: The casing of the used IoT device. Inside the casing, we can see a plateau meant to keep the circuit board aloft.

4.1.1.1 Tests

Tests were performed in various conditions. The experiments done on the device was done to determine the validity and accuracy of the distance sensor readings. As such, the amount of time spent on each test were not important, except during the experiment where the uptime of the device was to be determined.

A total of eight experiments were performed while the device was placed in different conditions. Firstly, an indoors test was done that verified the stability of the sensor data while in a stable environment. The experiment was done with a mostly clear and flat floor. This test was done during the weekend inside a common office for the *Open Distributed Systems* group at UiT. While some people may have been in the office at times, it was mostly empty.

Most of the remaining experiments were done with the device suspended over a snow-covered garden. In the second test, the device was outdoors while it rained lightly. The device was fastened to a plank during this deployment. In the following test, it rained heavily, and the device was fastened to a clothesline stand. A fourth test was done while the device was placed outdoors but inside a carport. This allowed the device to be affected by the temperature, while not being rained on. In this test, the device was placed over gravel. The fourth test was done right after the third, meaning that the device was still wet from the previous test. For the fifth test, the device was placed outside when it was overcast. In the sixth test, the device was placed outside in nice weather with only a thin snow layer beneath. The seventh test had the device placed over a shovelled part of the snow, while it snowed.

In the last test, the device was moved, and the conditions were changed artificially while the experiment progressed. All parts of the test were done in

(a) Plank deployment of the IoT device.

(b) Clothesline deployment of the IoT device.

Figure 4.5: Different deployment methods for the IoT device during the experiments.

nice weather. To start with, the device was placed over untouched snow. After a few hours, the snow underneath it was shovelled to make it flatter. Sometime later a few large buckets of water were poured over the snow. After applying the water, the snow was again re-shovelled and distributed evenly underneath the device. This test tested if the flatness of the snow played a large factor in the stability of the readings and checked how much the water content in the snow affected the distance readings.

4.1.2 Weather data

Weather data is collected from both self-deployed means and from outside sources. The intention behind this is to gather useful data relevant to people's willingness to travel outdoors. Since the weather close to the actual date of travel is the most likely to affect the user, the data collected is concentrated around the 12 hours before and after the current time, at the time of checking.

Data from the last 12 hours is collected from the Meteorological Institute (MET). The data is collected through the institutions *Frost* API. This data can be used to see if there is a high likelihood of weather caused road problems. Examples of these problems are mud, snow, and ice. These types of hindrances would take some time to form and some time to clear away. The 12 hours of knowledge could be used to analyse their probability of existence. The newest data point collected is the one closest to the current time. This measurement can be used to compare weather data readings from MET to other weather stations.

Only the latest value from the weather station at UiT — The Arctic University of Norway is collected in the current iteration of the project. This information is used to see how the locally measured weather can be used to enrich the readings collected from MET. The latest value from the *Frost* API can be thought

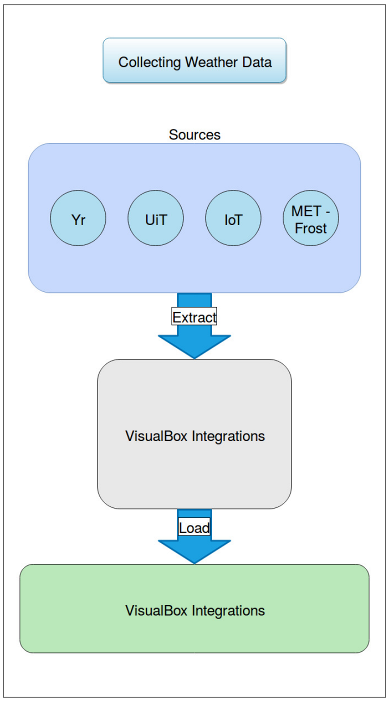


Figure 4.6: Flow of weather data for this project.

of as a *global* value for Tromsø, while the measurement from UiT is a *local* value. While the amount of information gotten from the difference between the two data values might be small, this might become more useful in places with tricky terrain. In such places, the terrain might shield some areas from weather conditions. This makes general readings from MET less accurate. A localized view of these spots can be used to create better-tailored advice to people living there.

The focus on short distance traveling for the project means that only information about the immediate past and future is needed. *Yr*, a group under MET, can give fine-grained forecasts close to the current date. Hourly forecast for about 12 hours into the future is collected. This was deemed enough to get an insight into what is believed to happen in a time frame relevant to traveling in an area for a day.

4.1.3 Crowdsourced

To use a third-party platform to communicate with users we need a way of tying message from them together. Twitter connects messages together through three main means: whom the poster is, specified *mentions*, and specified *hashtags*.

Tweets can only be posted by users with an account, providing them with some credibility by connecting an account to them. A user can add several *mentions* and *hashtags* to their posts to link them to a group of related content. A *hashtag* is a type of metadata tag used on social networks that allows other people to easily find messages with a specific theme or content [64]. *Mentions* are a type of hashtag that ties the content to a user instead of a specific theme. Both *mentions* and *hashtags* can be used as search parameters in the *Twitter Search API*.

As part of the project, a Twitter user was created to connect the messages posted to a common account. All users have a unique *mention* tag. This means that one can find posts relevant to the project by searching for the *mention* tag of the account in recent posts.

A JSON schema was used to bridge human and computer understanding of data. A JSON object contains a set of key-value pairs. A key being the descriptor of a field, while the value is the contents of the descriptor. An example of a key-value pair is that a key can be *first name*, while a corresponding value could be *John Doe*. Each Twitter message contains four key-value pairs, where the keys are: *street name*, *longitude*, *latitude*, and *message*.

Not all fields of the post are mandatory. A post must have a way of tying its content to a location. This means that either a *street name* or a *longitude and latitude* pair must be provided. The content of the post is in the *message* field. After the content is localized, it can be used to inform people about the situation at a specific site. An example of this would be if a user posted about a large amount of snow on a street. The government or a private party could then send out a snow plow to deal with the issue.

A widget was created with the purpose of helping users transform the information they wanted to post to a valid format. The widget is set up as a form with fields where the information one wants to declare can be easily stated. A set of possible message contents are given to help users provide data quickly. By limiting the options of message content, we can also define the information we are looking for. The message field can have anything as a value but providing users with predetermined choices helps them provide information quickly.

The Twitter APIs can be used by either an application or a user. An application can collect *tweets*, while a user can both collect and post them. The created widget takes the information received by the user and creates a link the user can use. This link moves the user over to Twitter, where the finished formatted *tweet* is ready to be posted.

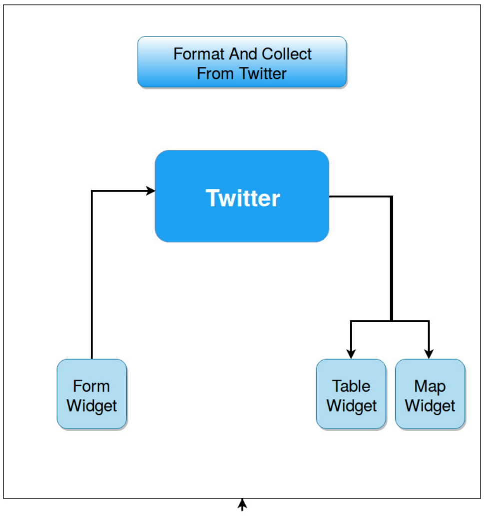


Figure 4.7: Flow of Twitter information.

In the Twitter collection integration, posts containing the specified *mention* is collected at a user given interval. This can be specified in a configuration page when installing the integration. One can also specify the number of messages in the configuration page. The posts are collected through the *Twitter Search API* [65]. The posts are mostly in a computer-readable format, but the *mention* tag needs to be stripped from the collected information. The remaining text can be parsed as a JSON object. After the created integration is installed into a VisualBox dashboard, any of the widgets available can collect information from it.

4.2 Presentation

The data collected from the VisualBox integrations are displayed through widgets in a dashboard. Multiple dashboards can be created, all using different integrations and widgets. This allows users to create dashboards with specific purposes. Only a single dashboard is used in this iteration of the project.

Five widgets are used in the created dashboard. Firstly, the common map that uses information collected from weather stations, deployed IoT devices, UiT, and Twitter posts. Secondly, a widget meant to help users format Twitter messages to a computer-readable format. Thirdly, a graph containing the temperature information from the weather data collected. Fourthly, a table over the last *tweets* posted. Lastly, a table of recommendations to users stemming from the weather data collected.

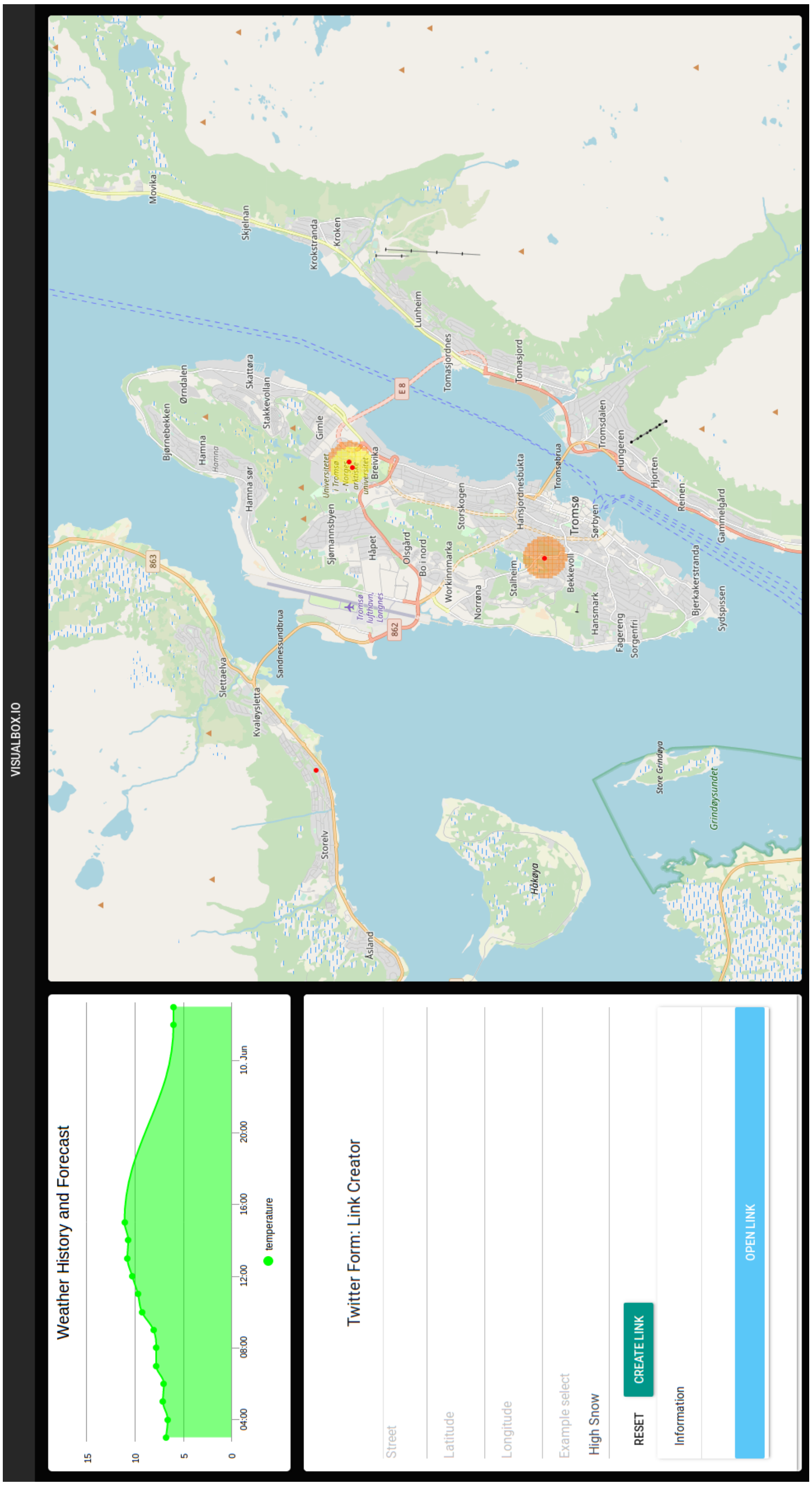


Figure 4.8: Part of the VisualBox Dashboard Created.

Widgets created were written in *JavaScript*, which is commonly used in web development. Libraries collected through Content Delivery Networks (CDNs) were used to allow complex widgets to be created. CDNs are geographically distributed networks of proxy servers and their data centres [66]. Simplified or minified versions of libraries collected through CDNs allows users to get complex web pages without taking up much space. In addition, the usage of CDNs decreases the need for storage of libraries by the servers and decreases the number of transmissions the servers need to do to clients.

Information collected to the common map is displayed in two ways: as transparent circles and opaque points. Data collected is represented on the map as points, showing where the information was collected from. By hovering over a point on the map, the relevant information from that point is displayed to the user. Temperature data is represented as transparent circles in addition to the points. The colour of the circles represents the temperature measured at a given point, with some colours representing warm areas and other representing cold areas. By using the map processing tool known as *GDAL*, the data was transformed into circles having an approximate radius of 1 km. The tool also uses a processing algorithm that fills the circles with their respective colours. In overlapping areas, the algorithm uses the colour of the average temperature of the two circles.

The weather widget used in this project was a graph. This graph widget was already created on the platform by the platform's creator, Pontus Edvard Aurdal. Temperature data collected from the historical weather data and the forecast is displayed in this graph.

The advice widget was created especially for the *RoadAhead* project. Weather data, both historical and forecast data, is used to create advice. The created advice is presented in three ways: as an *icon*, a *statement*, and a *reasoning*. An example of this is an icon of a raincoat, with the statement being that it will rain. The reasoning of this example could be that the forecast says that it is a likelihood of heavy rain in the near future. An icon and a statement are enough to give advice to a user, while the provided reasoning gathers trust in the statement. The icon is useful to experienced users, as a quick look can show the recommended outdoor wear of the day. A statement is useful for people who are unsure of what exactly the icon is supposed to represent.

In the current iteration of the widget, the advice is represented in a table. Three types of advice are given. The first regarding what type of footwear to use, while the others relate to what type of clothing to use. Footwear advice comes from the amount of downpour there has been or will be. Downpour information could indicate that there is mud or snow on the ground or that it will be soon. Clothing advice recommends the usage of either light or warm

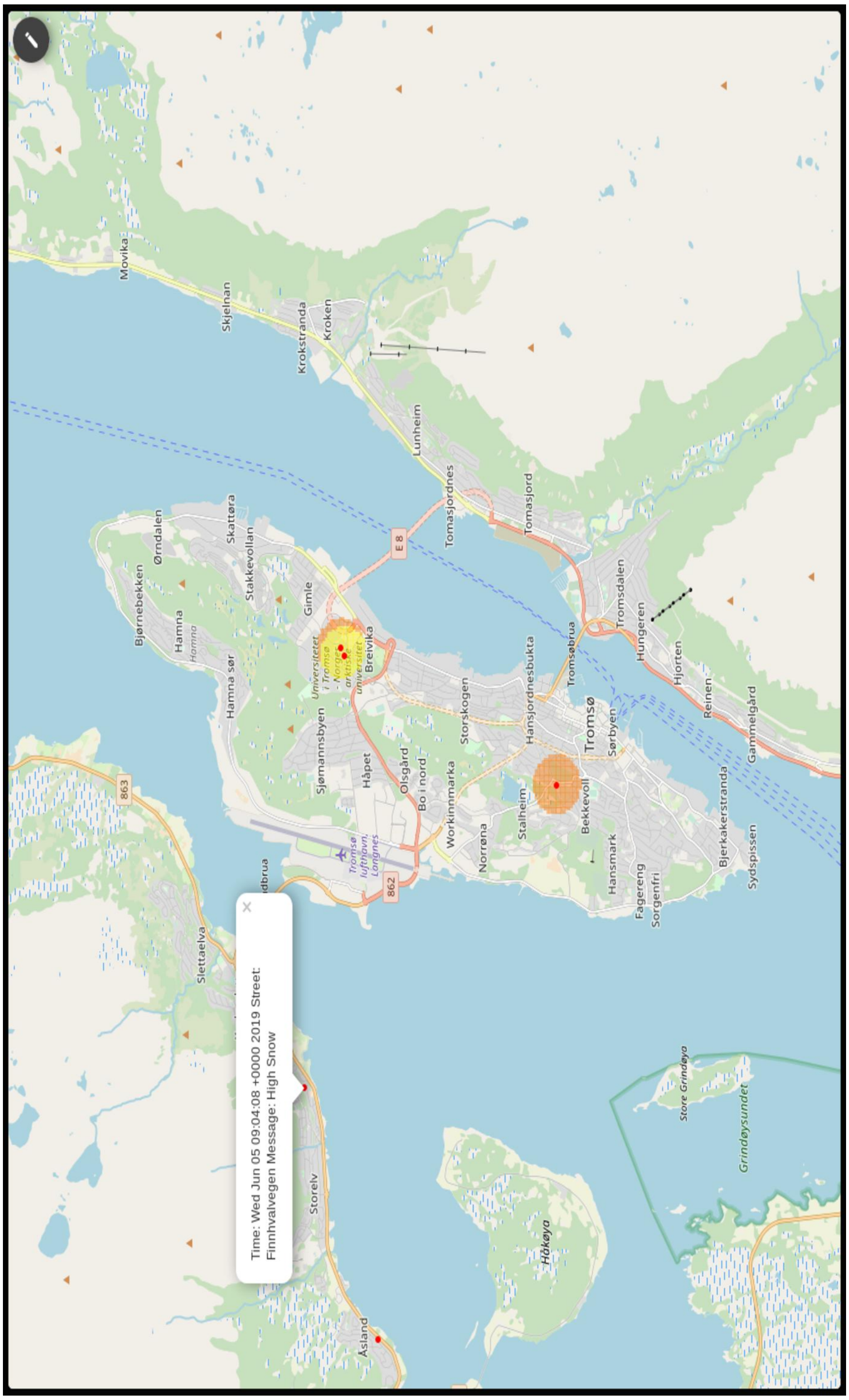


Figure 4.9: Common Map widget created to present information to users.

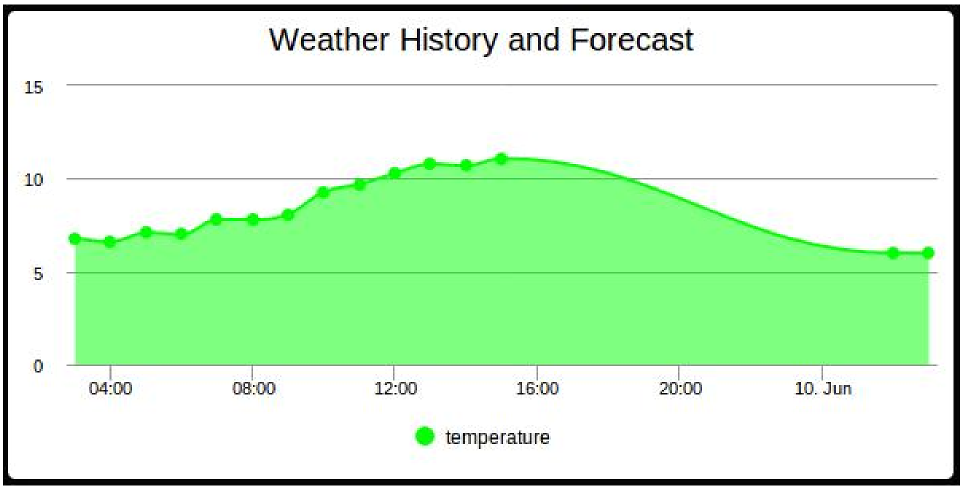


Figure 4.10: A graph widget containing temperature data from the Meteorological Institute in Norway.

clothing, and whether specialized clothing is recommended. The apparent temperature, calculated based on wind speed and actual temperature, affects the advice about wearing warm clothing [10]. *Specialized clothing* currently means either waterproof or windproof clothing and is affected by the expected downpour and wind speed in the coming hours.

The Twitter post formatting widget comes as a form stating the required fields to create a valid post. The message field is a drop-down selector with the following options: *high snow*, *patchy snow*, *ice*, *mud*, and *damaged road*. By using these descriptors, most commonly found and relevant issues can be reported easily. The second Twitter specific widget is a table containing the fields used in the Twitter formatting widget.

4.3 Summary

In this section, we looked at what was created during this project. Descriptions of the integrations created displays how different APIs are used and how information is analysed. The created IoT device and the experiments that were done upon it was looked at, together with how collection, analyses, and presentation of information is being done. In the following section, Evaluation, the results from the experiments done on the IoT device is looked at.

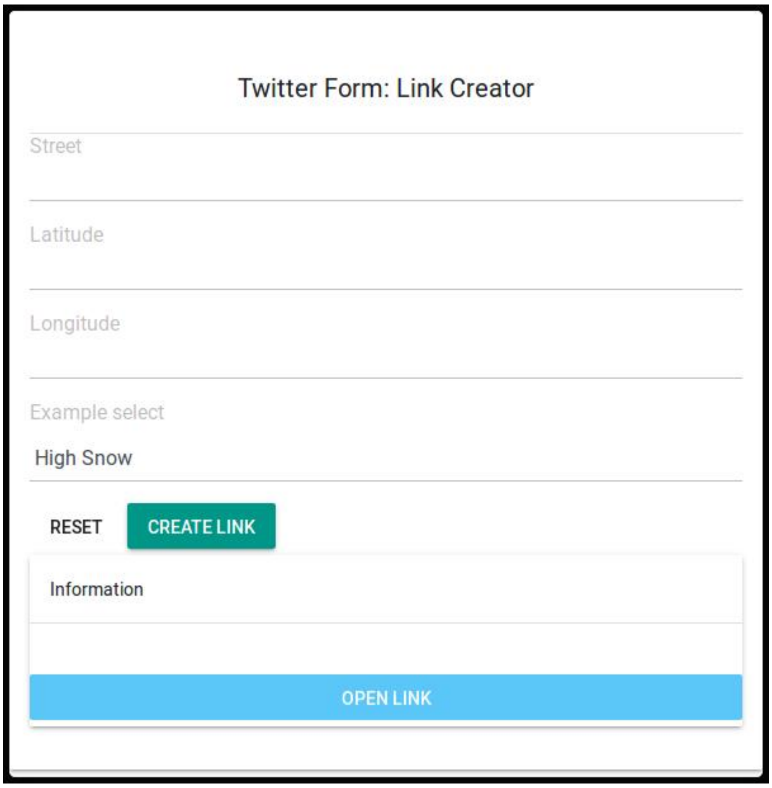


Figure 4.11: Twitter form widget created to help with post formatting.

Outdoors Clothing Advice

Icon↑	Recomendation	Reasoning
	Normal Shoewear	It has not rained/snowed a noticable amount in the near past, and it will not in the near future
	Light Clothing	It will not be or feel particularly cold in the near future
	Normal Clothing	There will not be heavy rain, snow or wind in the near future

Show entries 1 -3 of 3

Figure 4.12: Table showing clothing advice for the day.

"@road_ahead" Twitter Posts

Created↑	Street	Latitude	Longitude	Message
Sun Jun 09 14:01:05 +0000 2019	Åslandvegen	69.676995	18.770399	Damaged Road
Sun Jun 16 17:43:08 +0000 2019	Finnhvalvegen	69.686423	18.848217	High Snow

Rows per page: 10 1 -2 of 2

Figure 4.13: Twitter table widget created to present information from collected posts.

5

Evaluation

This section investigates the experiments done in this project. At the current iteration of the project, only experiments on the deployed IoT Device has been done. The results from these tests and some of the conclusions we can draw from them will affect how future designs of the device will change.

The device created has three main features that can be evaluated: the *uptime*, the *accuracy* of the measurements and the *physical design* of the device. Of the three features, the accuracy of the measurements is the most important. According to how accurate the device is, it can either measure actual snow height, differential snow height or it could be wholly unusable. The uptime of the device can be changed by exchanging batteries for more efficient or larger models and the physical design is meant to complement the measuring procedure.

Actual snow height is how much snow there actually is beneath the device. *Differential snow height* means that the device does not know how much snow there is beneath it, but over time it can measure how much the height has changed. This could be because the device has not been finely tuned, or because a patch of snow changes sonar readings with a set amount. An example of this is that a patch of snow could distort sonar readings to make the snow look 10 cm further away from the device than it is.

First, we will look at some of the logistics of the tests, before moving onto shortly explaining the test results. The eight experiments done with the deployed device will give us a look at its current uptime and accuracy. After explaining the results of the tests, we will investigate what conclusion we can draw from them and how to proceed in the future.

5.1 Logistics

The experiments were done by a device measuring the distance to a plane using sonar readings. Measurements done either by a ruler or measuring tape was done to have reference points to compare to the device readings. The closer the readings of the device are to the ruler readings the better.

A problem we must account for are measurement errors. We have two main types of errors, *statistical* ones and *systematic* ones. Systematic errors usually come from the measuring instruments. A statistical error, also called random errors, are caused by unpredictable changes in the experiment [67].

Measurements done by the sonar sensor was supposed to have a statistical error of around 1 cm when not connected to a temperature sensor. As the device has a beam pattern measurement area, errors could also occur if there are disturbances, like snowflakes, between the device and the measurement surface. Having a beam pattern for measurements of a surface means that it will likely work best on flat ones. A slope will either give measurement readings of the top distance, the bottom one or somewhere in-between those two. This could also not be constant. The effect of sonar measurement on snow could either be a statistical or systematic error on the measurements. If the error is not affected by the amount of snow or the water content in the snow, it will be a systematic error. Otherwise, it will be a statistical one.

Reference measurements for the distance between the device and the surface were done by either a ruler or a measuring tape. This means that we must take into human-error during the measurements and errors the measuring device itself brings. Taking consistent measurements by hand can be difficult, especially when measuring a malleable substance like snow. The ruler sometime sunk into the snow, and the snow was not always even. Depressions could also be on some parts of the snow. 2 cm was placed as the general statistical error on all ruler and measuring tape measurements. This could still be a bit on the low side due to the nature of the measured surfaces.

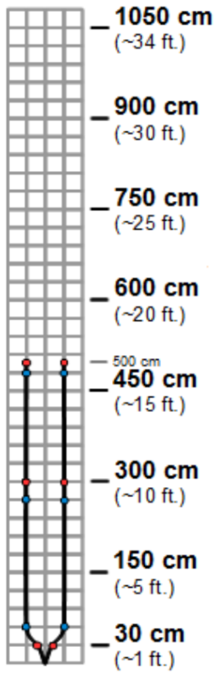


Figure 5.1: Beam pattern of MaxSonar sensor when measuring an 11-inch wide board [61].

5.2 Indoors Test

The first test to be looked at is the test done indoors. This test had the device fastened to an overhead beam at the *Open Distributed Systems* (ODS) laboratory at UiT — The Arctic University of Norway. As the test was done indoors, the device was undisturbed by the environment. Having people walk underneath the device could disturb the readings, so the test was done around the weekend. The device was deployed during a Friday and lasted until the Wednesday the following week.

The accuracy of the device can be verified from the test results seen during the weekend. In Figure 5.2 we can see that measurements taken during that time were mostly stable. The measured distance from the floor to the device was around 260 cm. After two days, the measurements started to become wildly inaccurate and the device started to only transmit data intermittently. This probably comes from the fact that the battery of the device started to become depleted. As IoT devices become low on power, they start to act unstable and unpredictable.

With the software at that time, the device uptime seems to be around 47 hours.

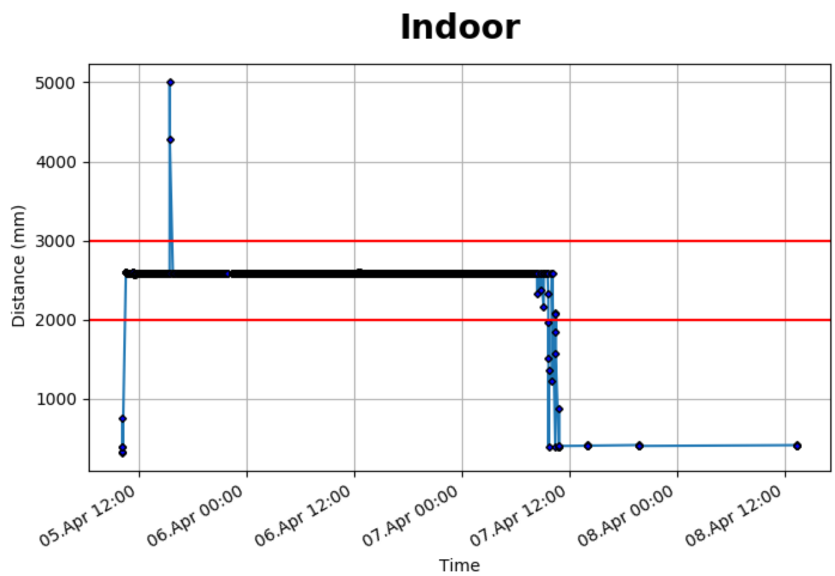


Figure 5.2: Measurements from the test done indoors on the IoT device.

Updates to the software have been done, so the actual uptime could be a few hours more or less than this value.

5.3 Good Weather Tests

The next two tests to be looked at was done during good weather. Both tests had the device mounted to a clothesline stick, which was suspended over a stand placed in a snow-covered garden. Snow completely covered the ground during both tests. In the first test, the snow remained undisturbed, while in the second the snow was modified twice. The modifications were done to test how different types of snow affects the sensor. This will be talked more about later in this subsection.

In Figure 5.3 we see the test results of the first test. During this experiment, the distance between the snow and the device was measured with a ruler 3 times. The first time was when the device was initially deployed. This was on the 24th of April, at 16:40 (24h CEST). At the time the distance was 121,4 cm. At 08:50 (24h CEST) the next day, the distance was around 121,6 cm and at 11:29 (24h CEST) the distance was measured to be 121,3 cm. While the measurements were not accurate, the stability shown can still be useful in measuring differential snow height.

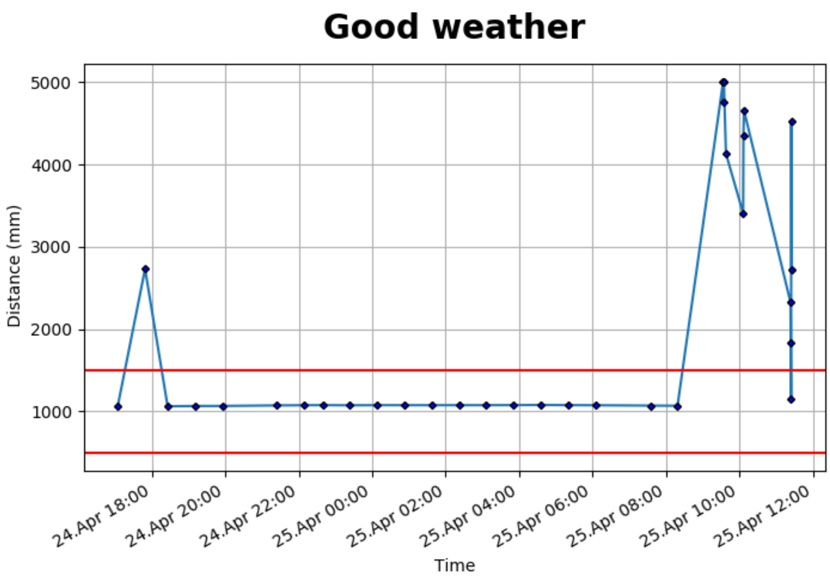


Figure 5.3: Measurements from the test done outdoors in good weather.

In the next test, the snow was modified twice while the measurements were done. Before the first modification, the device was placed over undisturbed snow. After some time had gone by, the first modification was implemented. The snow underneath the device was shovelled to get a more even surface. This test was done to look at how the accuracy of the device changed when measuring even snow compared to measuring normal, undisturbed snow.

After some more time went by, the second modification was done. This time, a few large buckets of water was poured on the snow, before the snow was shovelled anew to become even. This test was done to look at how the accuracy of the device changed depending on the water content of the snow.

The experiment was started on the 28th of April and ended on the 1st of May. Both modifications were done during the first evening. The test results can be found in Figure 5.4.

Reference measurements were taken 6 times with a ruler during the test. Of the 6 measurements, 5 were taken on the 28th of April. The first measurement was taken at 14:50 (24h CEST) when the device was first deployed. At that time the distance was 116 cm. Around 5 hours later, at 20:07 (24h CEST), the distance was around 118 cm. The first modification was done after the second measurement. After the modification, the distance was measured to have become 131,4 cm. At 22:33 (24h CEST) the distance was measured anew, this time reading 131,4 cm. The second modification was done at that time. After

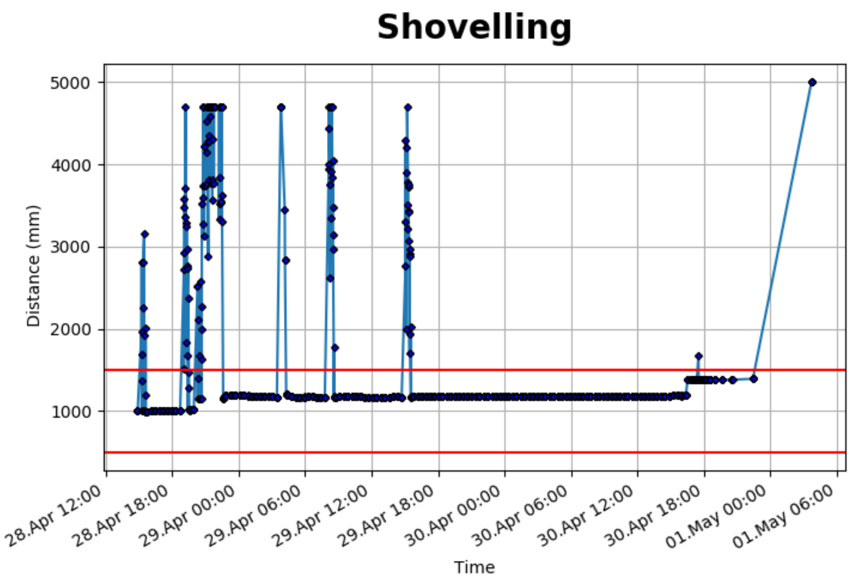


Figure 5.4: Measurements from the test with modification done to the snow.

this modification, the distance was read to be 138,3 cm. Three days later the distance was read again, this time having become 138,7 cm. This measurement was done at 16:01 (24h CEST) on the 1st of May.

The distance measured by the device differs from the ruler reading by about 16 cm at the start of the test. During the time before the first modification, the device seems to have stable readings in this area. After the first modification, the device measurements became highly inaccurate and unstable. The instability was reduced a few hours after the second modification. At that time, the difference between the ruler reading and the device readings was around 14 cm. Some conjectures to the reasoning behind this instability will be looked at later in this section.

5.4 Light Downpour Test

The test described in this subsection was done when it was raining lightly and quite windy. During this test, the device was mounted onto a plank and extended over a snow-covered garden.

This test started at around 12:11 (24h CEST) the 19th. of April and lasted until the device was collected at around 12:03 (24h CEST) the 20th. of April. At the time of deployment, the distance from the device to the snow was measured to

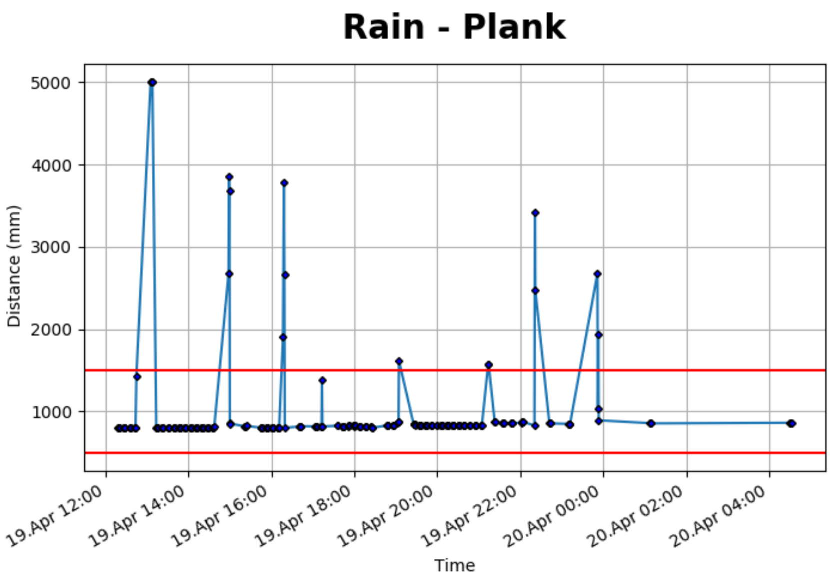


Figure 5.5: Measurements from the test done during light rain.

be around 77,5 cm. When the device was collected the distance was measured to be around 85,5 cm. The transmitted measurements from the device can be found in Figure 5.5.

During the experiment, the device had some trouble transmitting measurements. This can be seen in Figure 5.5, as the data points do not have a consistent distance between each point. In addition, there are no measurements between 04:48 (24h CEST) and the time of collection. The transmission problem has two likely explanations. Firstly, the bad weather and wind could have disturbed the transmissions. Secondly, the antenna of the device was pointing downwards during this test. Antennas can extend the range of transmissions by directing the signals and amplifying them. If the antenna directs the signal in the wrong direction, in this case downwards, the device will not be able to connect and transmit correctly.

During the test, the reference distance readings changed from 77,5 cm to 85,5 cm. This means that during the around 24 hours of deployment the snow height decreased by 8 cm. As snow height decreases, the distance increases. Measurements taken by the device started at around 79,6 cm and stopped at around 86,2 cm. This means that the device measured a change of 6,6 cm. As the device stopped transmitting early, the remaining difference could have happened during the time the device did not transmit data.

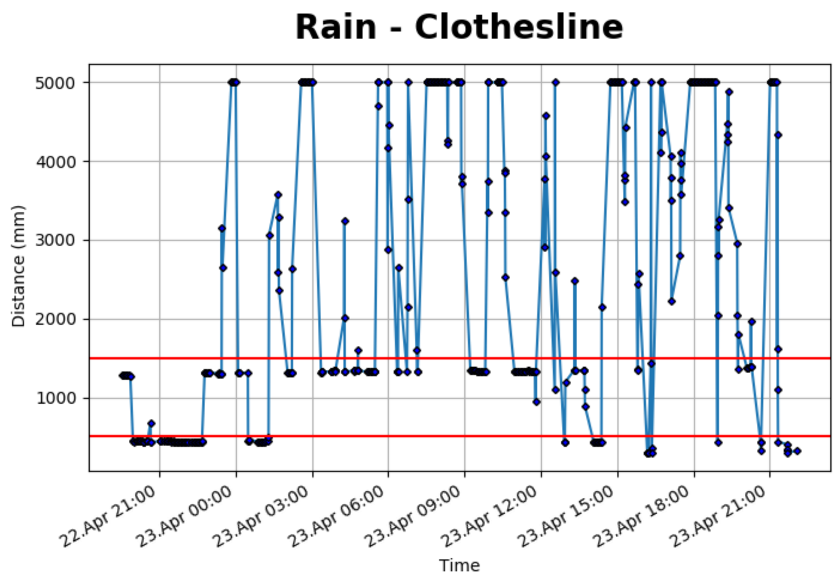


Figure 5.6: Measurements from the test done during heavy rain.

5.5 Downpour Tests

In the following two tests the device was affected by a heavy downpour of two different kinds. During both tests, the device was fastened to a clothesline and placed in a snow-covered garden. The first test was done during heavy rain, while the other was done while it snowed.

Figure 5.6 displays the resulting measurements transmitted by the device while it rained. The test was started at around 19:15 (24h CEST) on the 22nd. of April and ended the next day at 22:03 (24h CEST). Three reference measurements were done during this experiment. At the time of deployment, the measurement showed the distance to be around 129,0 cm. At 18:52 (24 CEST) on Tuesday the distance was around 134,0 cm. A final reading was taken when the device was collected. This reading showed a distance of 136,4 cm.

The distance readings from when it snowed can be found in Figure 5.7. The test was done from the 3rd of May to the 5th of May. Reference measurements were taken at the start and at the end of the test. The first reading showed a distance of 137,6 cm and the second a distance of 135,7 cm. This means that the snow height increased by around 2 cm during this period.

Results from both experiments show two similarities of note: instability and possibility. We can see from both tests that the measurements have values ranging from 300 mm to 5000 mm, the maximum and minimum values the

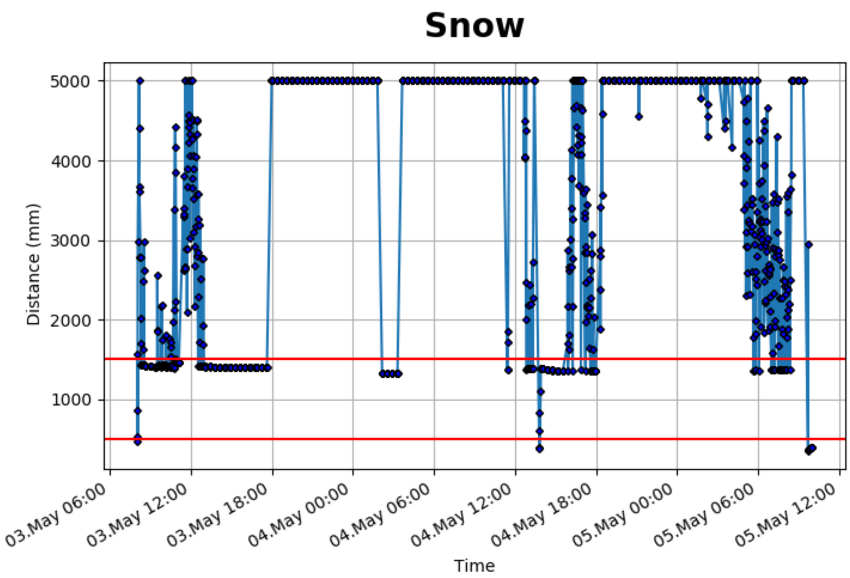


Figure 5.7: Measurements from the test done while it snowed.

device can measure. The other thing we can see is that there seems to be a line of semi-stable readings during the tests. We will investigate this and what it could mean for future designs later.

5.6 Wet Test

A second test was done immediately after the test with heavy rainfall. The device was moved inside a carport while the device was still wet from the rain. This test was supposed to look at the accuracy of the device when placed above an uneven surface. The results from this test could then be compared to tests taken over snow. This could be used to see if problems with measurements came from sensing an uneven surface instead of the material of the surface measured.

Results from this test can be found in Figure 5.8. During the test, the device was mounted to the clothesline with a distance of 147,0 cm from the gravel. From the graph, we can see that most of the readings from the device were around 100 cm away from the distance measure with a ruler.

The device was still wet from the previous test in addition to being placed over gravel. This could mean that the device was affected by the adhering water instead of the gravel surface. Adhering water droplets can cover the sensor

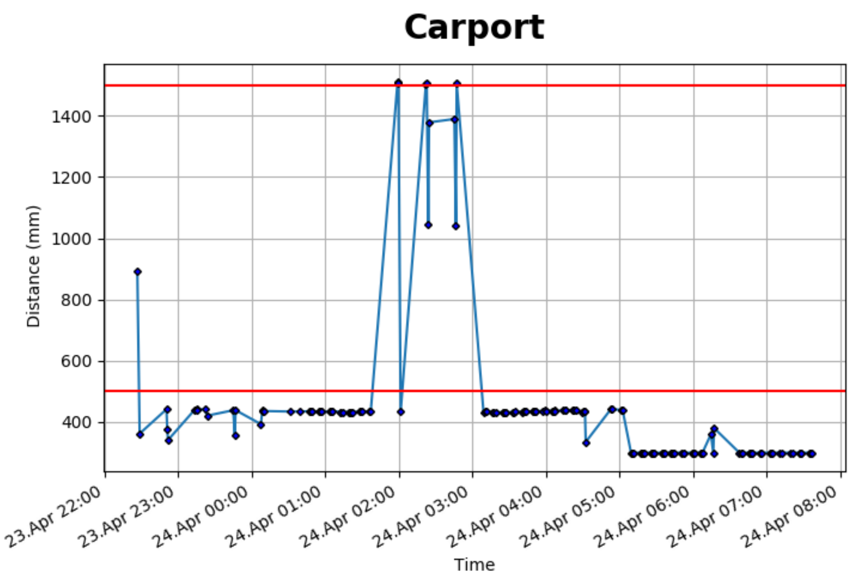


Figure 5.8: Measurements from the test done with the device inside a carport.

opening and cause the device to measure this, instead of the surface underneath. The possibility of a different casing design that will fix this problem will be looked at later in this section.

5.7 Thin Snow Test

The last test to be looked at was done in good weather over a small amount of snow. Figure 5.7 shows the results of the device readings. The test started around 08:29 (24h CEST) the 26th. of April and ended at 23:08 (24h CEST) the 27. of April.

During this test, the distance from the device to the snow was measured 4 times. At the time of deployment, the device was measured to have a distance of 125,0 cm to the device. At 15:49 (24h CEST) the same day, the distance was around 127,0 cm. Later, at 23:37 (24 CEST) on the 26th. of April, the distance was measured to be 127,6 cm. The final measurement was taken when the device was collected. This reading showed a distance of 133,7 cm.

From the graph in Figure 5.9 we can see that the device seems to have problems sensing a thin snow surface. In the beginning, the readings show great instability. While we can see some readings between 1000 mm and 2000 mm, many of the readings are either above or below those lines. What was

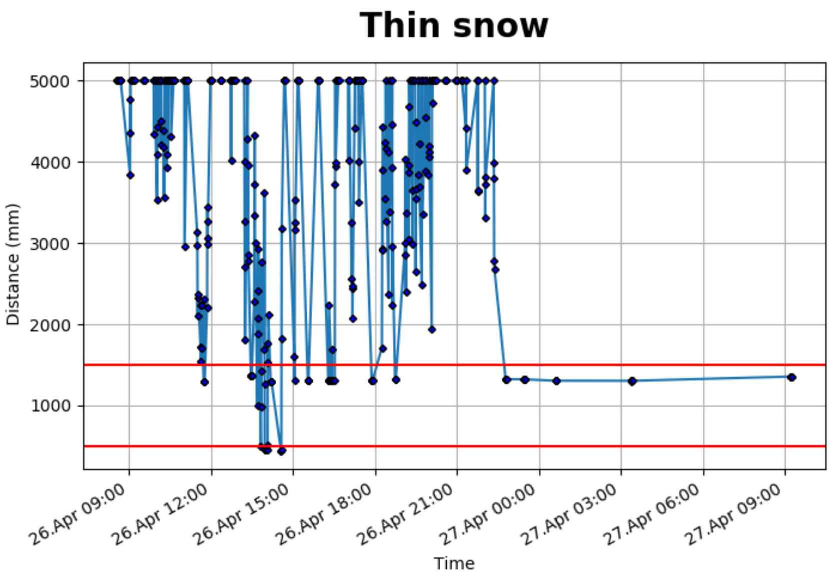


Figure 5.9: Measurements from the test done with the device over thin snow.

different in this test compared to the other tests done during good weather was that there was only a thin layer of snow beneath the sensor. The problem could be that the device measures both the snow and the wet ground underneath it. Another possibility is that thin snow absorbs sonar waves differently than thick snow. The actual reason for this must be tested in the future, as it was too little time to do it in this thesis.

Towards the end of the experiment, the device started to transmit data intermittently. During this test, the weather was good and there was not much wind. Together with the good weather, the antenna on the device was pointed upwards. Because of this, the reason for the lack of transmission is unknown. The problem could either be the device, or a random error that can be attributed to the LoRaWAN network.

5.8 Conjectures and Conclusions

In the following subsection, we will look a bit further into the results and what conclusions we can draw about the device. To get a better look at the measurements we will remove some of the readings that cannot be correct. This will show the benefit of introducing hard limits to readings, which will provide more accurate results. On readings over snow, we place the hard limits at 500 mm and 2000 mm. While the limits could be stricter, these limits could be used

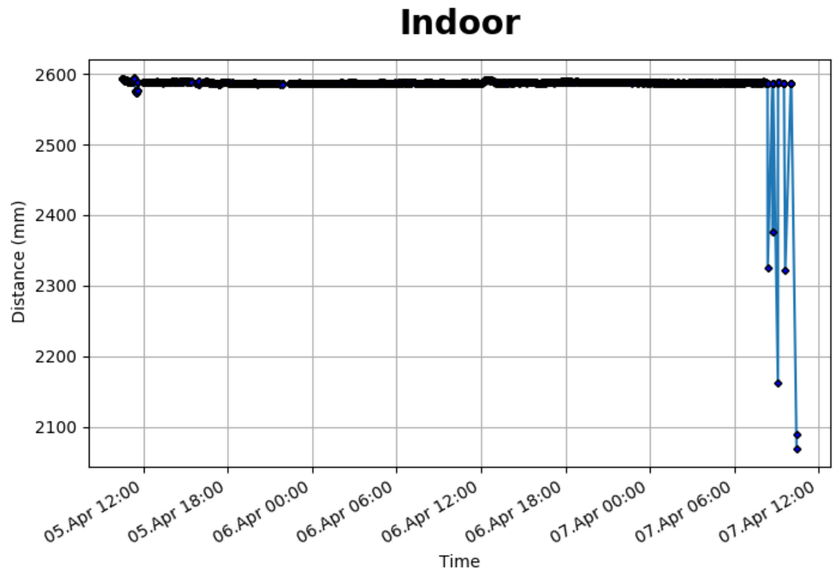


Figure 5.10: Indoor measurements with outliers removed. Limits set to 2000 and 3000 mm.

when the device is mounted higher than it was during these tests.

The good weather tests and the indoor test shows that the device can get stable readings close to the actual distance. In Figures 5.10, 5.11, and 5.12 we can see the relevant readings after removing those outside the hard limits. The already accurate results become even more accurate, with only a few outliers. While the readings done on the snow does show an inaccurate distance to the snow, the distance changes seem to be quite accurate. This means that even if the device cannot measure accurate snow height, it can be used to measure changes to it. The readings from the device can then be used with an initial reading as a reference height.

There are two current conjectures to the cause of the instability in the test with snow modifications. The first conjecture is about the consistency of snow at different depths. If snow becomes significantly more packed together compared to snow closer to the surface, it could affect readings in strange ways. This reasoning seems unlikely, as the readings seem to have become stable after some time. The second conjecture has to do with particle snow. As the snow was disturbed by the shovelling, uneven layers of powdered snow were distributed on top of the measured surface. Powdered snow could have been the reason for the unstable readings as it may have deflected the sonar waves or given responses at different times to the same sonar wave. After some time, the powder either became part of the other snow or got blown away by the

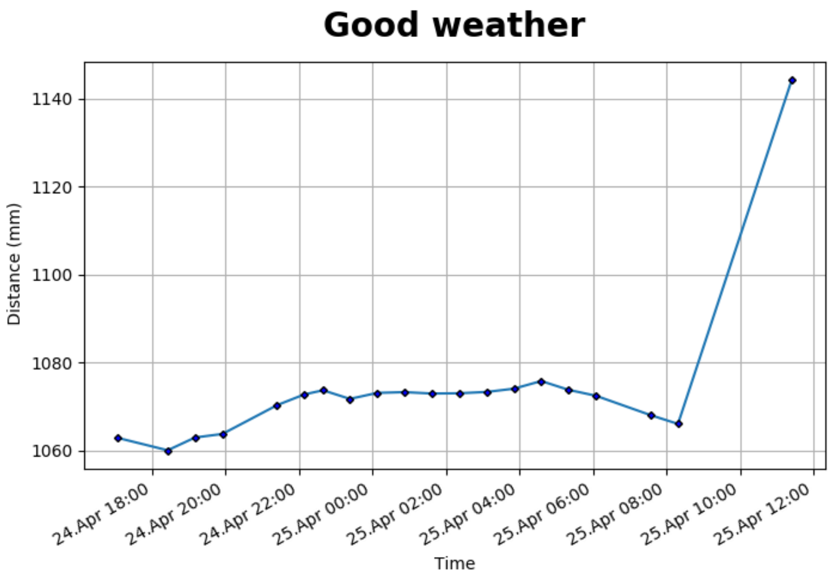


Figure 5.11: Good Weather measurements with outliers removed. Limits set to 500 and 2000 mm.

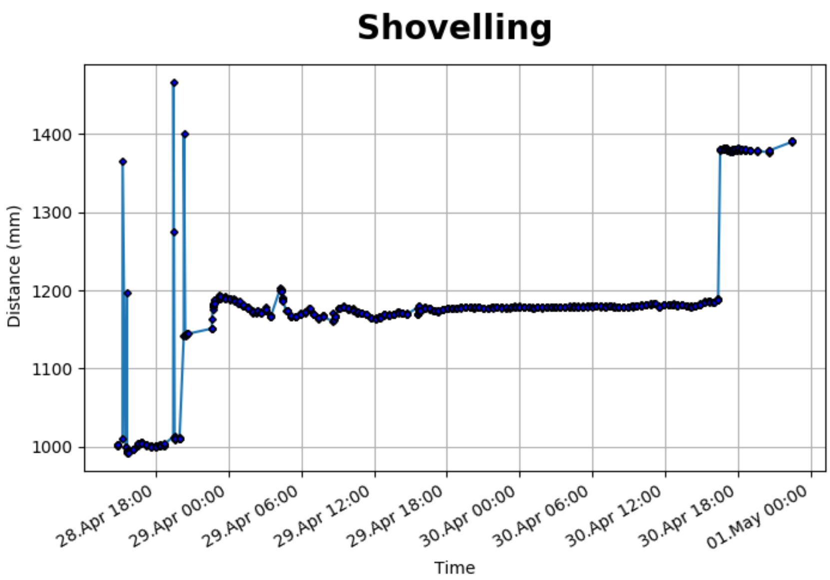


Figure 5.12: Snow Modification measurements with outliers removed. Limits set to 500 and 2000 mm.

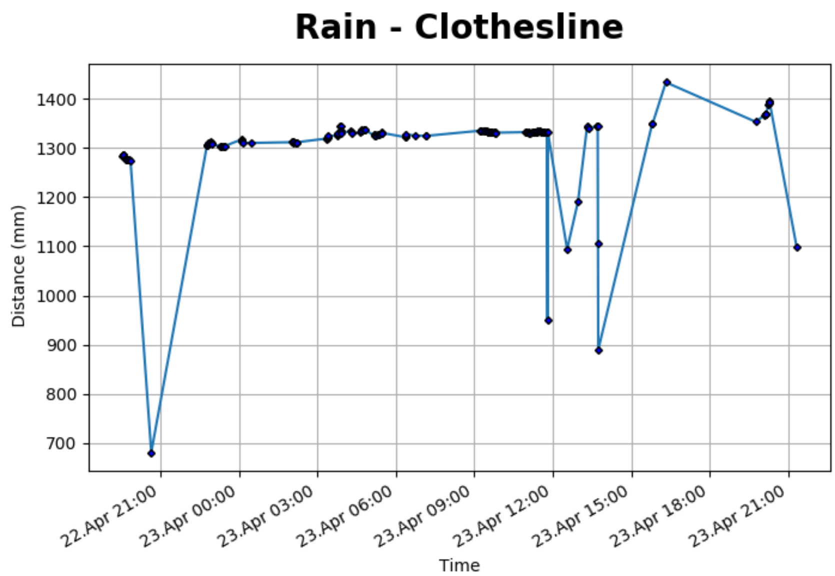


Figure 5.13: Heavy Rain measurements with outliers removed. Limits set to 500 and 2000 mm.

wind. This is the reason the distance measurements became stable after some time.

From the measurement tests taken by the device, there are a lot of tests that show highly unstable readings. When placing in the hard limits on the tests we can see that some of the remaining readings fall closer to the expected values. Still, the remaining readings show that downpour will limit the usage of the device. While readings over thin snow could be a problem, the device could still be useful in places where a hard packed snow layer is present. An example of this is roads during the winter or ski tracks.

In the carport test, the remaining results are useless. The problem is that most of the values are inaccurate, which means that having a wet device is a fatal problem. This problem could be mitigated if we design the casing in a way that removes the water droplets from adhering to and around the sensor. In Figure 5.16b a possible design for this purpose is shown. The design would make the lowest point of the device lower than the sonar opening. By having the opening have a conical design the water droplets are moved away from the opening. In addition, the conical design will not interfere with the beam pattern of the sensor. The new design was not implemented because of the time constraints on this thesis but will be investigated in the future if the project continues.

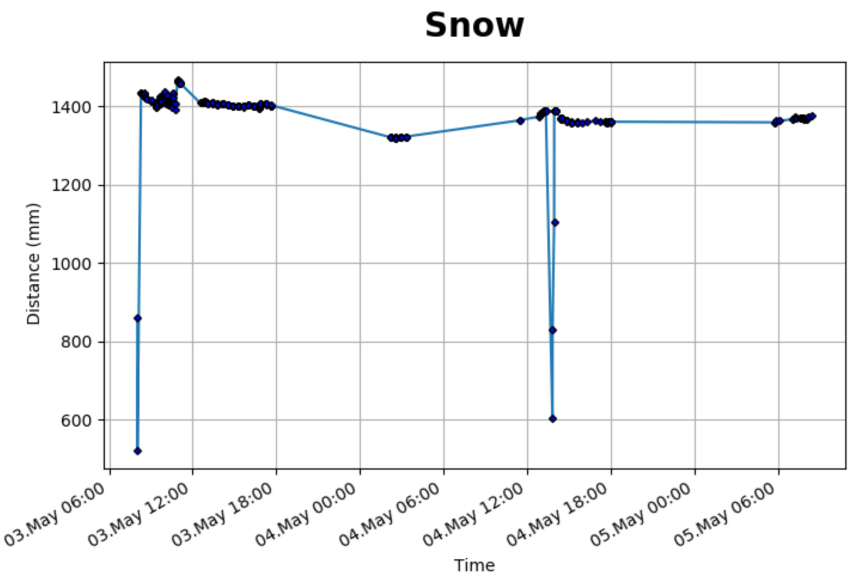


Figure 5.14: Snowing measurements with outliers removed. Limits set to 500 and 2000 mm.

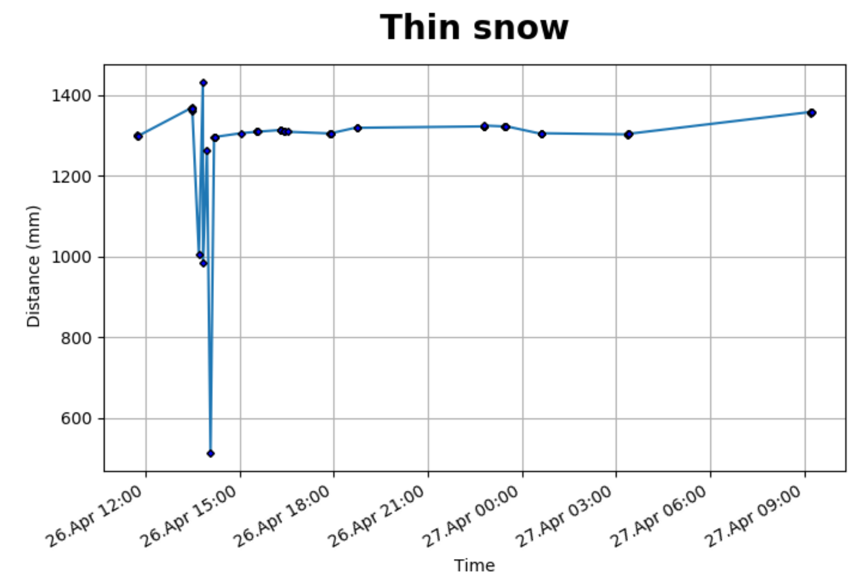
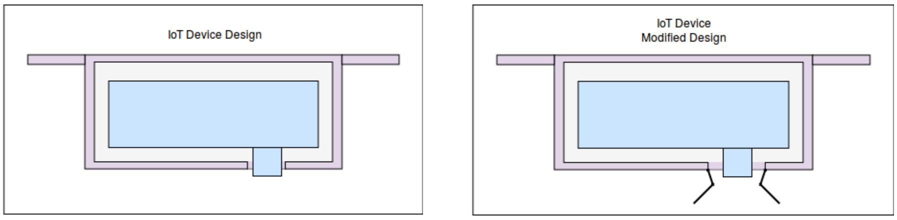


Figure 5.15: Thin Snow measurements with outliers removed. Limits set to 500 and 2000 mm.



(a) Original IoT device design. **(b)** Modified IoT device design.

Figure 5.16: The original and modified IoT device designs.

5.9 Summary

In this section, we looked at the experiments done on the deployed IoT device. From the initial testing, we can see that the device could work to detect changes in snow height when the weather is good. The device seems to have problems with downpours, which is likely because of the detection of particles between the sensor and the sensing surface. This can be mitigated by using hard limits to remove impossible or improbable readings, or to enter a deep sleep mode when a lot of impossible readings are found. In this way, we will be able to conserve power and can continue normal measurements when the downpour ends.

In the following section, we will look critically at what has been done so far. Together with a look into what has been done, we will investigate possible ways to proceed with this project.

6

Discussion

In this section, the overall progress of the project will be looked at. The steps from project creation to the state of the project so far will be discussed. Afterward, what could have been done differently will be looked at. When looking at what could have been done differently, other people's solution to similar problems will also be looked at.

6.1 Project progress

The focus of this project is to help people travel with certainty. Information was collected, analysed and presented to provide users with useful information to this end. As the collected data comes from multiple sources, a data warehouse was the solution selected to fulfil the needs of the project. While other applications, like the transportation recommender application from the Related Work section, does something similar, there are some fundamental differences. These will be addressed in some of the later subsections.

6.1.1 Selecting Focus

Five information types and three data sources were identified to be of relevance to this project. Information about *accidents*, *parking*, *people*, *roads*, and the *weather* could be of use. These information types could either be collected by us, from third-party sources or from crowdsourcing.

First-party data sources are used to supplement data not found in outside sources. By using third-party sources as the main sources of information, the number of needed self-deployment units decreases. This allows the project to focus on its main goal of providing analysed data to users. Because of time limitations, the project was focused on collecting information from all three data sources. By collecting from these data sources, a foundation for future research could be created.

Weather data and information about the roads was selected to be the starting focus to collect information about. Information about the weather affects all outdoor types of travel, an example being that mist can create problems for both cyclists and drivers. Road information is in this project defined as information about the ground people and vehicles travel upon. This includes both highways and tracks.

The data sources selected was based on the information types we wanted to collect. To collect weather data, MET was selected to be the third-party data source. Additionally, weather data was collected from UiT — The Arctic University of Norway as the first-party data source. Information about the roads was collected both from users over Twitter and from UiT. From this, the goal of setting up data collection from the three data source types was completed.

The data sources were researched to see what types of data were available and what could be created to enhance the data. This research gave us a starting point to proceed from. Historical weather data and weather forecasts could be collected from MET. UiT has its own weather station but had few data sources concerning the roads. Small messages containing information can be posted to Twitter, but these messages do not have a set schema.

6.1.2 Collection and Analysis

As the focus was to decrease uncertainty in travel, information about the immediate past and future was the focus of information. Weather data from 12 hours into the past and into the future was collected from MET for this reason. To start the work of tailoring information, weather data from UiT was collected to supplement the data from MET. In this way, data from MET could tell us something about the general weather conditions in Tromsø, while UiT could give a more localized view around the university.

To use Twitter as an effective communication platform we needed to process the posts to produce computer-readable messages. The focus of looking into Twitter became to find out how to format messages easily, and how to tie all message created for this project together in an easy to collect manner. This was achieved through a formatting widget and the usage of metadata tags added to the messages.

Collecting information about the roads using Twitter requires users to actively report the state of the roads, so an additional way of collecting information in a passive manner was researched. From the research, a snow height sensing IoT device was developed and tested. By connecting the IoT device to the LoRaWAN network, live data collection from outdoors became possible.

The collected data from the different sources were used to create different displays showing the outdoor conditions in Tromsø. Weather data was displayed through coloured circles on maps, displaying places with higher or lower temperature compared to a trusted value. The trusted value, in this case, was the data collected from MET. Values collected from UiT and from the deployed IoT device was used to give localized values to the map. Data from the deployed IoT device and the crowdsourced messages were also displayed on the map, as dots that could be interacted with. By hovering over a dot, information about that location was displayed. An example of this information is the temperature measurements taken there.

Some data was displayed using different types of tables. The weather data collected from MET was analysed and used to display recommended outdoor clothing if one wanted to travel outdoors. The apparent temperature was calculated using wind speed and the measured temperature. This allows us to create advice based on how cold a human feels it is outside. The information collected Twitter was also displayed using a table, which should allow a user to easily see information posted by other people. As the information displayed contains the location concerning the message, users can use this table to find messages concerning themselves. This could be done by looking at the street name of the given information.

6.1.3 Evaluation of work done

The work done can be evaluated in their component parts, or as a complete whole. In this subsection, the component parts will be discussed, while the complete whole will be discussed later in the Discussion section.

By looking at the measurements of the deployed IoT device we could see that the current way of detecting snow height is not sufficient. While we got stable readings from most of the readings in good weather, bad weather affected the readings in major ways. Heavy rain affected the deployed IoT device in a way that stopped it from collecting correct readings even after the rain had stopped. This could be mitigated with a different case design that protects the sensor from adhering water droplets. This will be investigated in the future if the project continues.

The readings done in good weather shows that the device could work if the sensor is only going to measure how much the snow height changes over time. Problems occur when the snow is not even or when the snow is thin. Deployment of the device to places where this does not occur could be one way to avoid this. An example of this could be to place the sensor in areas where there is already a thick layer of snow. As similar IoT devices are being researched, and in some places used [27], more research can be done into the usage of sonar to measure snow height. Usage of another sensor type, or multiple sensors, could be investigated to increase accuracy.

The current collection of weather data seems to be enough for the current state of the project. Collection of weather data 12 hours into the past and future provides enough data to do short-term analyses of the outdoor conditions. Additional data could be collected to better measure snow height changes over the winter, but different terrains would make the resolution of analyses low. This will be looked at later in the Discussion section.

MET and UiT are both trustworthy deliverers of accurate weather data. By arranging the trustworthiness of the accuracy of the data by their sources we can display both general and localized weather situations. Displaying this on a map allows the data to be easily absorbed by users. While a graph displaying the temperature in the close past and imminent future might be unnecessary, the information provides some use to users. The graph displays how the weather is probably going to change and the general trend of changes in a close time frame. This allows users to make their own decisions and shows them some of the data used to create advice given to them.

Using Twitter as a crowdsource communication platform is functioning properly. Twitter already has ways of tying messages together and tying messages to users who posted them. Through APIs, these posts and their relations can be extracted. As the platform also has a way of allowing the creation of finished posts through links, the amount of work the users must do to format messages is low.

The created formatting widget allows users to easily create messages, while the map displays the information posted by users in an easy to understand manner. Providing the data as a table in addition to displaying it on a map allows users to directly look at information posted by users. This could make it easier for users to find out the general situation of a place by looking at several messages at the same time. While the current way of collecting information is acceptable for general, non-sensitive data, a more secure and private measure must be created for more sensitive and personal information. Tying data to people would require users to either have their own profile or storing all their data in a standalone software. Possibilities of these will be looked at later in the Discussion section.

6.2 Additional and Alternative methods

The information collected and the analyses done upon them in this project is only a subset of the possible ones that can be explored. In this subsection, more ways to extract and analyse information will be looked at. Before going into detail about collection and analysis, the overall project can be looked at and discussed.

Collecting, analysing and presenting information about how it is outdoors and giving travel advice is the focus of this project. This is something that has been done before by others. An example of this is weather forecasts on the television. Another is the travel planner of *Google Maps* which can give travel options based on your mode of transportation. Most of these applications are narrow or specialized. To again use *Google Maps* as an example, they only collect information about traffic conditions, while ignoring the weather. A project that can be compared to the current project, in both goal and procedure, is the recommender application introduced in the Related Work section [20].

The recommender application was a digital nudging project that also wanted people to use environmentally friendly transportation options. Information was collected on both the users and the environment. This information was used to personalize nudges with relevant data. While this project is similar in these ways, the application was a bit shallower than this project aims to be. The recommender application collects information about the weather, the user and the travel distances. This project wants to look at several more points of reference that might impact traveling. Information about accidents, road conditions and a user's health are only some of the information points this project will eventually investigate. While the recommender application was also not finished, the goal seemed to be set a bit lower than this project. As such, the application cannot be a complete substitute for this project.

This project was compared to the recommender application because they have similar goals and procedure methods. The problem is that the application was a bit shallower than this project. There are many applications that do similar things to this project but have narrower goals. The focus of promoting equipment and betterment in travel through healthier and green transportation is also a place where this project remarks itself. This also something that has been done before. Another example of a similar application is the *Tromsø Mobility App*. This application is another project from the NUDGE group at UiT. The application is meant to nudge people to choose greener transportation options [68, 69]. What the project described in this thesis will contribute with is a broad way of looking at different outdoor and personal conditions to help people be better informed before traveling.

6.2.1 General Architecture

The current architecture uses a data warehouse approach as its general design. This approach has three components: *sources*, a *processing unit*, and an *endpoint*. In the source component are MET, UiT and users. VisualBox integrations take care of the processing unit's *Extract, Transform and Load* work. After processing, the information is loaded into VisualBox widgets to be presented to the users. The widgets are the current endpoints of this project. While the design works as intended, a future design could add more components in-between the current ones.

Between the sources and the processing unit, pre-processing components can be added. Some pre-processing is already done on the data, an example being that the deployed IoT device already filters and average its values to make it more accurate. An example of a component that could be added is a pre-processing unit that takes in the information from weather stations hosted by both MET and UiT and processes the data to a common format.

Middleware can be added between the processing unit and the endpoint. This could be done to filter out unnecessary information or to distribute information to users. An example of a component to add is middleware meant to work as a publish-subscribe system. This is similar to what is explained in Andersen et al. [12] and showed in Figure 2.1. In this case, the *Sense* part of the system would be the *sources* component, the *Analyse* part would be the *processing unit* and the *Inform and Nudge* part would be the *endpoint*. A difference would be that the publish-subscribe capabilities would be added as an additional part between the endpoint and the processing unit.

6.2.2 Collection and Analysis

Alternative or additional methods of collecting, analysing and presenting data needs to be done with the background of providing useful information. Collect-

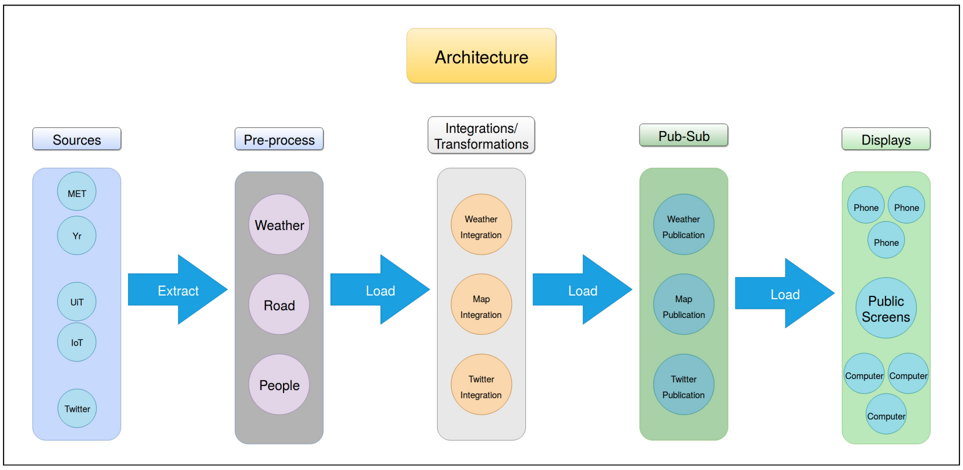


Figure 6.1: A modified version of the project architecture.

ing information from sources should be done in a way that provides the most useful data, while not including non-relevant ones. In this subsection, some of the alternatives available will be looked at.

6.2.2.1 Weather Data

One of the reasons weather data is collected from multiple sources is that we can create local and global views of the weather conditions. These views are based on the data trustworthiness and the presence of sensors. The number of weather stations from the current sources are limited in some ways. While MET has weather stations throughout Norway, they are spread out. By collecting from more stations, we can create a fine-grained and fault tolerant system weather map. Instead of deploying multiple weather stations ourselves, this can be crowdsourced to smaller, private weather stations. This weather data can be compared to the larger weather stations around Norway to look at the probability of correctness.

NETATMO is a provider of small, private weather stations and has APIs that can be used to collect weather data from those. The company has multiple weather stations throughout the world, which can add a lot of data to our weather map. Collecting information from a large amount of weather station can create a more seamless weather map. This might also provide more information than needed, but superfluous data can be removed.

Collecting weather data from other organizations is also a possibility, like collecting information from NASA weather stations. A widespread weather station network can allow for fine-grained weather maps to be created with

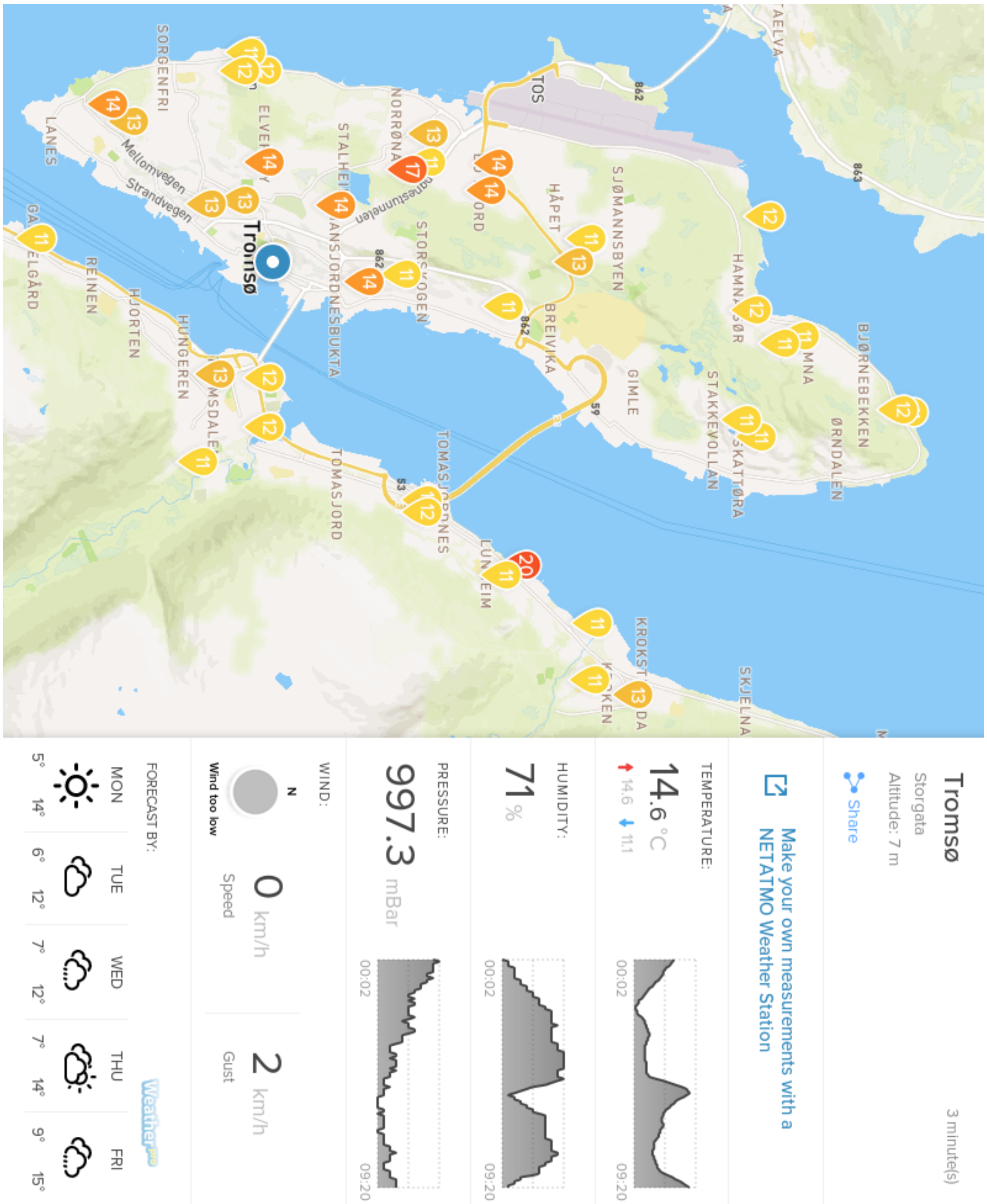


Figure 6.2: Weather Map with information from NETATMO weather stations.

highly accurate data. The necessity of this amount of detail is not certain, as there is little danger of large differences in weather over small areas.

Additional weather data and data about its effect on the environment can supplement the data already collected. Different camera types can be used to collect data about mud and water content in the soil or snow. Reflection can be used to detect gatherings of water or ice patches. Specialized soil sensors can find the water content in the soil. Readings about this type of data can be used to give advice about the recommended clothing for trips. The amount of water in snow, or the amount of snow in general, could be used to analyse the probability of avalanches or similar disasters.

6.2.2.2 Road Data

The currently collected road information in this project comes from the deployed IoT device and from Twitter. Additional data about the roads can be inferred from weather data. An example of this is that the advice integration looks at the amount of downpour in the last 12 hours to give footwear recommendations. A large amount of rain in the last 12 hours could mean that mud, sludge or snow is likely on the ground. This would warrant footwear that can protect feet from this type of terrain.

The deployed IoT device collects data about the snow height by measuring the distance between the snow and the sensor. Snow height could be inferred from measured snowfall, but this would fail to account for snow melting and terrain differences. Weather data from MET can contain measured snow height at the weather station, but snow height is not the same in all places. Slopes would have deeper snow than hills because of gravity and the amount of sun that hits the snow.

Snow detection can be done by camera imagery. Usage of colour or reflection are some ways to detect snow and ice. In the Related Work section, multiple ways to detect water content in soil is mentioned, and some can additionally be used to detect snow.

The crowdsourced road data could contain any location-specific data that fits within the Twitter size limit. To get more complex and easier to use data an application that can help users to format such messages is needed. Placing more metadata in the forms can help with refining the data to an easier to create and read format. An example of this is that users can say that the information transmitted is about road damage, which could then be refined to be information about potholes. Starting out by saying that the information is about road damage would remove some of the possible messages one can select to send, like a message about high snow.

Additional data can be collected from other crowdsourced sources, like Facebook or from a standalone application created specifically for this purpose. *Gatami* is a website that people living in Tromsø can use to report road problems. Information from such sources would be like what can be collected from Twitter.

More information about roads can be collected to supplement or enrich the data currently collected about roads. Specialized systems, like the ones provided by the *Observator* group or specialized vehicles like the *Roadstar*, can collect a lot of data about roads. By using these types of solutions, information about slipperiness, road damage, and the presence of ice can be collected. While specialized systems like these might be expensive, smaller and cheaper solutions are also available. Systems like the *RoadMonitor* and the *Pothole Patrol* are examples of this.

By providing simple sensors that can be connected to normal vehicles a wide detection area can be made. By placing sensors on buses or taxis, multiple routes can be monitored without obstructing the traffic any more than usually done. Solutions like *SODiCS* can be used in areas where it is difficult to place sensors and server stations. By transmitting data from one vehicle to another, users can get updates about the road ahead of them that are newly collected.

Monitoring traffic also falls under monitoring roads. Information about the amount of traffic on vehicular and non-vehicular roads can be used to infer which roads are well maintained, damaged or congested. This means that traffic data can both enrich the current data collected and bring with it its own useful information.

The simplest way to collect data about traffic in Norway might be to use Floating Car Data (FCD), as there are many people with smartphones in developed countries. This type of data can be collected from Google and other data providers. Another way of collecting FCD would be to equip buses or taxis with GPS units. By transmitting their position, we can infer the speed they have and how much congestion there is at their location. This unit could also be connected with accelerometers and gyroscopes like in the *RoadMonitor* system.

The bus system in Tromsø currently provides timely updates about how late buses are on their respective routes. This could be used in a similar manner to discover where congestion or road problems are. Some static approaches to collecting traffic data, like image recognition, can also be used if such data can be easily collected. More specialized static approaches, like laser tripwires or magnetic readings, can be more expensive and time-consuming to deploy. The amount of detail one can get from such systems is likely higher than needed, as the exact number of cars or car types on a road is not directly useful to this project.

6.2.2.3 Parking

Collecting information about parking spaces can useful by itself and to enrich data about the roads. Systems to detect cars exist, like the ones provided by *Windmill Software* [46]. Such systems can be used to count cars and by connecting them to a common collection centre, the number of parking spaces used in an area can be found.

Data about parking spaces can enrich the data about roads and traffic congestion by inferring where a lot of people are traveling. By providing information about where there are parking spaces left, the amount of idle traffic can be decreased as search time for such spaces decreases. If the information is provided in a standardized format, multiple sources and application can use it.

Finding different parking areas in Norway can be done through an API provided by *Statens Vegvesen* [70], the organization maintaining the public roads in Norway. The API for parking registration can be used to extract information about parking areas and parking providers [71].

6.2.2.4 People

Information about people can be used to personalize advice given to users. Collecting information from smartwatches and health applications can help provide data that is useful to specific people. The health of a person can affect what types of non-vehicular options are available, like bicycling or walking.

The *Twitter Search API* used to collect information only collects data from the last 7 days. While non-permanent information does not need long term storage, some permanent or semi-permanent information does. Examples of this are information on places containing air spread allergens and information about the allergies of a person.

The information posted to Twitter is collectible through APIs and through the Twitter website. Any person, even without an account, can find the information posted there. Privacy-sensitive data would need to be posted or transmitted through other means than Twitter to preserve a client's rights. The creation of a standalone software might be needed in the future for this reason.

Collecting information about a user's travel options, like the availability of cars or skis, can affect what types of travel advice is applicable to a person. Providing information about allergies can change travel routes, like circumventing parks during pollen seasons. How different advice affects different people can also be used to refine the advice given to them. An example of this could be how formal the advice given should be, or if a person only accepts nudges in the afternoon.

6.2.2.5 Accidents

Accidents and mistakes add uncertainty and danger to traveling that are hard to fully handle. This is something that can never be fully removed but can be worked on continuously to be reduced. While accidents on roads can be hard to predict, their precursors can be mapped out. An example would be to report places with especially slippery road sections. These sorts of problems can easily be fitted into the current reporting system using Twitter, as it takes in a location and a message.

When accidents happen, as it certainly will, reporting such issues needs to be handled in a way that reduces the overall problems presented. An example of this could be to replan travel routes as traffic accidents are reported. Replanning travel routes based on sudden obstacles to already established routes can be done in more settings, like when road maintenance reports come in.

Removing stress from drivers can reduce the number of traffic accidents. While accidents can be reported by users through crowdsourced measures, established sources of accident information can also be extracted from. An example of such a source in Tromsø is the organization which maintains the public roads, *Statens Vegvesen* and the police. Both entities have Twitter accounts that can report such issues. While these accounts do not give out computer-readable information, the data they do produce can be analysed for useful information.

Statens Vegvesen also has its own APIs where one can collect traffic messages through *Datex* [72]. This API can transmit real-time data about weather, travel times, camera imagery and traffic messages that contain information about road maintenance and driving conditions. *Datex* is a European standard for exchanging traffic data.

By integrating this API into the project, together with using crowdsourcing to collect accident reports, most road problems can be found. Information about specific roads can be found through the Datex API, which means that roads of an established travel plan can be checked against it. If obstacles are found, another travel path can be created that circumvents these issues.

6.2.3 Visualization

Information is currently presented to users through the VisualBox website. The visualization tool allows users to collect, analyse, and present data to outside sources. The current widgets are written in *JavaScript*, *HTML* and *CSS*. Third-party libraries are collected and used from Content Delivery Networks (CDN).

The current widgets, while functional, are not the most aesthetically pleasing. This will be improved upon as the project continues. A common template needs to be found to improve consistency, and the information can be stated in clearer ways. Personalization of presentation will also become possible as the project moves forward. As an example of an area needing improvement is that there is only a visual way to perceive the data presented, making the information unusable by visually impaired people. This could be partly resolved by adding website reader capabilities to the widgets or by creating it in a standalone application.

As mentioned in the previous subsections, the number of message options provided by the Twitter post formatting widget can be expanded upon. A dynamic solution, transforming the form based on previous choices can be investigated to create an easy to use widget. An example of this would be that a type field in the format can change the options for the message content. This would allow for a shorter list of more suitable options. In this way, the choice of a *road damage* type message can give options like *pothole damage* and *cracks in the road* as the message content.

The widget presenting raw weather data about from 12 hours into the past and future might show some useless data. This is because it shows so much information about the future without doing any analyses on it. A shorter list of previous data could be enough. On the other hand, by allowing this data to be shown users can see why certain advice was given from the advice widget. The advice widget only shows human-readable advice and reasoning, while hiding the data that brought its conclusions. By using a widget to display data to users directly, the dashboard provides users with openness about the data.

While the tool works well as a temporary measure in its current iteration, there are some limitations. Currently, only one integration can be used by a widget at a time. This means that either large integrations are needed, or the same code needs to be reused in multiple integrations. As the complexity of integrations increases, specialized widgets that can read specific data collections are needed.

A standalone application, either hosted on the internet or on phones, is a likely step forward. This could allow for a greater degree of privacy and personalization. Phone applications allow for widgets to be created, something that could allow users to only look at the information they find relevant. Another possibility could be that the application can schedule a trip based on personal needs and weather conditions, and add it to a user's calendar in a seamless manner. By using a local installation on a user's phone or computer, storage of personal data can be done on devices trusted by the user.

6.3 Summary

The main goal of this project was to provide people with information that could reduce the uncertainty they have during travels. This was done with the thought that uncertain terms and variables during travels make people default to a safer mode of transportation, despite the environmental or health effects that will bring.

The work that was done during this project created the beginning of a working data warehouse that could fulfil this project's goals. Data from multiple outside and first-party sources is collected to the data warehouse and is used to provide a cohesive, situational awareness of the outdoor conditions. Currently, not all the possible information is collected. In addition, the analyses done is still in its beginning phases. In this section we described some of the work that needs to be done in the future. In the next section, Conclusion, we will go through what we can take from this project so far. In addition, we list some of the future work that will be done next if this project continues.

7

Conclusion

The goal of the *RoadAhead* project was to lessen the uncertainty people have about traveling outdoors. By increasing a user's knowledge, their confidence brought by certainty should allow them to make bolder travel choices. Instead of going with the safe choice of a private car, the user might choose to use a bicycle if he knew the good weather would continue throughout the day. Information about where snow is allows users to choose skiing tracks to their destinations. The work in this project has been about collecting data, analysing it and then presenting it to users.

Information gathered about the weather and the roads were analysed and presented to users to help inform them about the situation outside. This information was gathered from outside sources, like the Meteorological Institute (MET), from people, and from self-hosted solutions. As the project progressed, the amount of information available showed that there was a lot that could be done in the future. By acquiring more knowledge, a glint of the amount of 3rd order knowledge out there became apparent [11].

Weather data was collected from MET and UiT — The Arctic University of Norway. MET and UiT both host weather stations. Additional information was also supplemented by a deployed IoT device. A map was created to show how the temperature changed in a local area. In addition to this map, recommendations about outdoor clothing were created from analyses done upon the weather data. This provided the foundations of the program planned at the outset of this project.

Information gathered from people and the self-deployed IoT device provided additional data to the created map. Layering information upon a map will allow users to easily find data relevant by looking at what was happening close to their location. While the only a small amount of information is currently collected, this too creates a part of the foundation needed to complete the project in the future.

Creating IoT devices and connecting them to systems can be convoluted. Research into what others have done and the development process of this system is similar to what is described in *The Five Orders of Ignorance* [11]. Collecting information, trying to find the solutions to the devices program, and testing it allows the knowledge of what to do to emerge. By connecting a sensor to the created data warehouse, a foundation for information extraction from sensor networks was established.

While there are similar projects to this one, this project does bring green nudging towards a slightly different direction. The persuasive recommender application [20] meant to nudge towards environmentally friendly transportation is an example of a similar project. The main difference between the recommender project and the goal of the *RoadAhead* project is that the *RoadAhead* project tries to aim broader and deeper. The current project brings focus to looking broader and specializing information to specific users to a higher degree. An example of this is that the *Road Ahead* project will eventually start looking into health conditions and integration with other applications. Such applications could be a smartphones pedometer program. Looking at several aspects that affect traveling, collecting from multiple data sources, and using this information to give recommendations is the main contribution of this project. It can be looked at as a project that collects information to provide information.

While the foundation of the project has been established, a lot more can still be done. The data collected has been focused on only two of the five identified information types from the Introduction section. Information about people, accidents, and parking still needs to be collected. Data on people will allow us to personalize advice, which can reduce the amount of unneeded advice given to a user. Information on parking can reduce idle driving which comes from looking for a vacant parking lot. Accident information can provide real-time relevant data like travel route changes caused by traffic accidents. Of the information types collected from, a lot more can still be collected or analysed. Detection of potholes and slippery roads are some examples of this. Analysing the probable snow height from the downpour information and other weather data is another possible enhancement. In addition to collecting and analysing more data, personalization needs to be investigated. To do this effectively a user profile setup needs to be looked at.

7.1 Future Work

As the progress of knowledge acquisition shows us how much knowledge we lack, so will the amount of work to be done increase as the current work progresses. It might be that this knowledge-acquisition project will require years to create a fully established application that encompasses everything initially wanted. Personal advice that nudges a user's decisions about traveling outdoors, while considering his best interests. This is also to be done while editing the suggestions made based on a user's wants and their access to modes of transportation. And after finishing a smaller version, bringing this project out into the world on a large scale.

As the project launches on a larger scale, it will also have to consider how different countries behave. Examples on what to consider are the available sources of information in a country and how the culture affects how nudges work on the people there. In this section, we will list some of the work that is currently seen to be needed in the future.

7.1.1 Remaining data types

The remaining three information types need to be collected from. Some of the benefits of the different information types have been explained repeatedly. As such, we will not go into everyone again here. The amount of information that can be collected can be seen from the Discussion and Related Work sections, where different collection methods and projects are looked at. Of the three remaining information types introduced in the Introduction section, personal information might be the most important. This is because talking to people and giving them the information they as individuals need, can help them in making the choices they want.

Personal data can be collected as user profiles or through options given at the start of an application. Information like the health of a person and his age can affect the advice given. An example of this is, suggesting warmer clothes at higher temperatures than normal if a person is sick.

7.1.2 User system

To collect data about users a certain amount of privacy and security is needed. The creation of specialized accounts can help with both issues. Allowing users to have their own account allows us to configure what they see to only be what they find useful. Information stored about users will be tied to a user profile, and all the user information can be stored together securely. This makes it so only the user herself can see the information she provides about her person.

7.1.3 Standalone Application

The current use of VisualBox has some issues that could be resolved by creating a standalone application for this project. While the tool works well as a development tool, the lack of dashboard sharing makes it so that every user of the project needs to set up their own informational pages. A standalone application would allow users to have a standardized view, that could be modified to accommodate the wants of a user. This would simplify the customization process. The current way of doing things would also require users to create their own accounts on different sites, like for the *Frost* API of the MET and for the *Telenor Managed IoT Cloud* service. This is not something we can easily assume all users are willing to do and would act as a deterrent to users.

Bibliography

[1] Waze Mobile. Waze. `https://www.waze.com/no/`. [Online; accessed 4-June-2019].

[2] Statistisk Sentralbyrå. Statistisk Sentralbyrå. `https://www.ssb.no/transport-og-reiseliv/faktaside/bil-og-transport#blokk-1`. [Online; accessed 20-June-2019].

[3] Anne Ellaway, Sally Macintyre, Rosemary Hiscock, and Ade Kearns. In the driving seat: psychosocial benefits from private motor vehicle transport compared to public transport. *Transportation Research Part F: Traffic Psychology and Behaviour*, 6(3):217 – 231, 2003.

[4] Linda STEG. CAN PUBLIC TRANSPORT COMPETE WITH THE PRIVATE CAR? *IATSS Research*, 27(2):27 – 35, 2003.

[5] Donald J. Wuebbles and Atul K. Jain. Concerns about climate change and the role of fossil fuel use. *Fuel Processing Technology*, 71(1):99 – 119, 2001. Fuel science in the year 2000: Where do we stand and where do we go from here?

[6] Marilena Kampa and Elias Castanas. Human health effects of air pollution. *Environmental Pollution*, 151(2):362 – 367, 2008. Proceedings of the 4th International Workshop on Biomonitoring of Atmospheric Pollution (With Emphasis on Trace Elements).

[7] World Health Organization. Mortality and burden of disease from ambient air pollution. `https://www.who.int/gho/phe/outdoor_air_pollution/burden/en/`. [Online; accessed 20-June-2019].

[8] Marko Tainio, Audrey J. de Nazelle, Thomas Götschi, Sonja Kahlmeier, David Rojas-Rueda, Mark J. Nieuwenhuijsen, Thiago Hérick de Sá, Paul Kelly, and James Woodcock. Can air pollution negate the health benefits of cycling and walking? *Preventive Medicine*, 87:233 – 236, 2016.

[9] Natalie Mueller, David Rojas-Rueda, Tom Cole-Hunter, Audrey de Nazelle, Evi Dons, Regine Gerike, Thomas Götschi, Luc Int Panis, Sonja Kahlmeier, and Mark Nieuwenhuijsen. Health impact assessment of active transportation: A systematic review. *Preventive Medicine*, 76:103 – 114, 2015.

[10] Wikipedia contributors. Apparent temperature — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Apparent_temperature&oldid=889930133`, 2019. [Online; accessed 4-June-2019].

[11] Phillip Armour. The Five Orders of Ignorance. *Communications of the ACM*, 43, 10 2000.

[12] Anders Andersen, Randi Karlsen, and Weihai Yu. Green transportation choices with IoT and smart nudging. In Muthucumaru Maheswaran and Elarbi Badidi, editors, *Handbook of Smart Cities: Software Services and Cyber Infrastructure*, pages 329–352. Springer, 2018.

[13] Richard Thaler and C Sunstein. *NUDGE: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press (Publisher), 2008.

[14] Randi Karlsen and Anders Andersen. Recommendations with a Nudge. *Technologies*, 7(2), June 2019.

[15] Lucia Reisch and C Sunstein. Do Europeans Like Nudges? *Judgment and decision making*, 11:310–325, 08 2016.

[16] C Sunstein. Do People Like Nudges? `http://nrs.harvard.edu/urn-3:HUL.InstRepos:16147874`, 05 2015.

[17] Gabriele Esposito, Penélope Rojas, René Van Bavel, and José Vila. Nudging to prevent the purchase of incompatible digital products online: An experimental study. *PLOS ONE*, 12:e0173333, 03 2017.

[18] Matthias Lehner, Oksana Mont, and Eva Heiskanen. Nudging – A promising tool for sustainable consumption behaviour? *Journal of Cleaner Production*, 134:166 – 177, 2016. Special Volume: Transitions to Sustainable Consumption and Production in Cities.

[19] Markus Weinmann, Christoph Schneider, and Jan vom Brocke. Digital Nudging. *Business & Information Systems Engineering*, 58(6):433–436, Dec 2016.

[20] E. Bothos, D. Apostolou, and G. Mentzas. A recommender for persuasive messages in route planning applications. In *2016 7th International Con-*

ference on Information, Intelligence, Systems Applications (IISA), pages 1–5, July 2016.

[21] Hendrik Bruns, Elena Kantorowicz-Reznichenko, Katharina Klement, Marijane Luistro Jonsson, and Bilel Rahali. Can Nudges Be Transparent and Yet Effective? *Journal of Economic Psychology*, 65:41–59, 01 2018.

[22] The Norwegian Meteorological Institute. About the Norwegian Meteorological Institute. https://www.met.no/en/About-us/About-MET-Norway, 02 2018. [Online; accessed 4-June-2019].

[23] Wikipedia contributors. NASA — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=NASA&oldid=899616321, 2019. [Online; accessed 4-June-2019].

[24] Gavin A. Schmidt, Robert B. Schmunk. GISS Surface Temperature Analysis. https://data.giss.nasa.gov/gistemp/stdata/, 03 2019. [Online; accessed 4-June-2019].

[25] NETATMO Weather Station. https://www.netatmo.com/no-no/weather/weatherstation. [Online; accessed 4-June-2019].

[26] NETATMO Weather Map. https://weathermap.netatmo.com/. [Online; accessed 4-June-2019].

[27] UNIS The University Centre in Svalbard. Better avalanche warnings in and around Longyearbyen. https://www.unis.no/better-avalanche-warnings-in-and-around-longyearbyen/. [Online; accessed 4-June-2019].

[28] Wikipedia contributors. Soil moisture sensor — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Soil_moisture_sensor&oldid=895045005, 2019. [Online; accessed 4-June-2019].

[29] Brian Leib, Jay Jabro, and Gary R. Matthews. Field Evaluation and Performance Comparison of Soil Moisture Sensors. *Soil Science*, 168:396–408, 06 2003.

[30] Edwin T. Engman and Narinder Chauhan. Status of microwave soil moisture measurements with remote sensing. *Remote Sensing of Environment*, 51(1):189 – 198, 1995. Remote Sensing of Land Surface for Studies of Global Chage.

[31] A L Rankin and L H Matthies. Passive sensor evaluation for unmanned ground vehicle mud detection. *Journal of field robotics.*, 27(4):473–490, 2010.

[32] Arturo L. Rankin, Larry H. Matthies. Multi-Sensor Mud Detection. `https://www.techbriefs.com/component/content/article/tb/techbriefs/physical-sciences/6554`, 01 2010. [Online; accessed 4-June-2019].

[33] AIT AUSTRIAN INSTITUTE OF TECHNOLOGY. Road Condition Monitoring. `https://www.ait.ac.at/en/solutions/road-condition-monitoring/`. [Online; accessed 4-June-2019].

[34] Adham Mohamed, Mohamed Mostafa M. Fouad, Esraa Elhariri, Nashwa El-Bendary, Hossam M. Zawbaa, Mohamed Tahoun, and Aboul Ella Hassanien. Roadmonitor: An intelligent road surface condition monitoring system. In D. Filev, J. Jabłkowski, J. Kacprzyk, M. Krawczak, I. Popchev, L. Rutkowski, V. Sgurev, E. Sotirova, P. Szynkarczyk, and S. Zadrozny, editors, *Intelligent Systems'2014*, pages 377–387, Cham, 2015. Springer International Publishing.

[35] Jakob Eriksson, Lewis Girod, Bret Hull, Ryan Newton, Samuel Madden, and Hari Balakrishnan. The Pothole Patrol: Using a Mobile Sensor Network for Road Surface Monitoring. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*, MobiSys '08, pages 29–39, New York, NY, USA, 2008. ACM.

[36] Kasun De Zoysa, Chamath Keppitiyagama, Gihan P. Seneviratne, and W. W. A. T. Shihan. A Public Transport System Based Sensor Network for Road Surface Condition Monitoring. In *Proceedings of the 2007 Workshop on Networked Systems for Developing Regions*, NSDR '07, pages 9:1–9:6, New York, NY, USA, 2007. ACM.

[37] A. Ghose, P. Biswas, C. Bhaumik, M. Sharma, A. Pal, and A. Jha. Road condition monitoring and alert application: Using in-vehicle Smartphone as Internet-connected sensor. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops*, pages 489–491, March 2012.

[38] K. Ito, G. Hirakawa, Y. Arai, and Y. Shibata. A Road Condition Monitoring System Using Various Sensor Data in Challenged Communication Network Environment. In *2015 IEEE 29th International Conference on Advanced Information Networking and Applications Workshops*, pages 518–523, March 2015.

[39] Tian Lei, Abduallah A. Mohamed, and Christian Claudel. An IMU-based traffic and road condition monitoring system. *HardwareX*, 4:e00045, 2018.

[40] Ashish Dhar. Traffic and Road Condition Monitoring System. Technical report, Indian Institute of Technology, 2008.

[41] Guillaume Leduc. Road traffic data: Collection methods and applications. *Working Papers on Energy, Transport and Climate Change*, 1:1 – 55, 01 2008.

[42] Andrey Nikishaev. Tutorial: Making Road Traffic Counting App based on Computer Vision and OpenCV. https://medium.com/machine-learning-world/tutorial-making-road-traffic-counting-app-based-on-computer-vision-and-opencv-166937911660. [Online; accessed 4-June-2019].

[43] Wikipedia contributors. Google Maps — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Google_Maps&oldid=900107209, 2019. [Online; accessed 4-June-2019].

[44] Tromskortet. Troms Travel planner. https://www.tromskortet.no/?lang=en_GB. [Online; accessed 4-June-2019].

[45] Entur. About Entur. https://www.entur.org/about-entur/. [Online; accessed 4-June-2019].

[46] Windmill Software Ltd. Vehicle Sensing: Ten Technologies to Measure Traffic. https://www.windmill.co.uk/vehicle-sensing.html. [Online; accessed 9-June-2019].

[47] Tromsø parkering. Tromsø parkering. https://tromso-parkering.no/. [Online; accessed 4-June-2019].

[48] Tromsø kommune. GATAMI. https://www.gatami.no/. [Online; accessed 4-June-2019].

[49] Pontus Edvard Aurdal. VisualBox – A Generic Data Integration and Visualization Tool. Master's thesis, UiT The Arctic University of Norway, 2019.

[50] Pontus Edvard Aurdal. What is VisualBox? https://docs.visualbox.io/. [Online; accessed 20-June-2019].

[51] Space-Track.org. Space-Track. https://www.space-track.org/auth/login. [Online; accessed 4-June-2019].

[52] LoRa Alliance Technical Marketing Workgroup 1.0. LoRa Alliance: A technical overview of LoRa and LoRaWAN. Technical report, LoRa Alliance, November 2015.

[53] Meteorologisk Institutt. WeatherAPI. `https://api.met.no/`. [Online; accessed 4-June-2019].

[54] Meteorologisk Institutt. What is Frost. `https://frost.met.no/index2.html`. [Online; accessed 4-June-2019].

[55] UiT THE ARCTIC UNIVERSITY OF NORWAY. Weather observations, about. `http://weather.cs.uit.no/about/`. [Online; accessed 4-June-2019].

[56] yr.no. Free weather data service from Yr. `https://hjelp.yr.no/hc/en-us/articles/360001940793-Free-weather-data-service-from-Yr`. [Online; accessed 4-June-2019].

[57] Wikipedia contributors. Semantic Web — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Semantic_Web&oldid=898556668`, 2019. [Online; accessed 4-June-2019].

[58] telenor connexion. Managed IoT Cloud. `https://www.telenorconnexion.com/managed-iot-cloud/`. [Online; accessed 4-June-2019].

[59] pycom. LoPy4. `https://pycom.io/product/lopy4/`. [Online; accessed 4-June-2019].

[60] pycom. Pysense. `https://pycom.io/product/pysense/`. [Online; accessed 4-June-2019].

[61] MaxBotix Inc. *HRLV-MaxSonar®-EZ™Series*. [Online; accessed 10-June-2019].

[62] Arne Munch-Ellingsen. LoRaWAN coverage. private communication, May 2019. [Online; accessed 8-May-2019].

[63] Wikipedia contributors. Interquartile range — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Interquartile_range&oldid=898338894`, 2019. [Online; accessed 4-June-2019].

[64] Wikipedia contributors. Hashtag — Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Hashtag&oldid=896669899`, 2019. [Online; accessed 4-June-2019].

[65] Twitter. Search Tweets — Standard search API. https://developer.twitter.com/en/docs/tweets/search/api-reference/get-search-tweets.html. [Online; accessed 4-June-2019].

[66] Wikipedia contributors. Content delivery network — Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/w/index.php?title=Content_delivery_network&oldid=897105609, 2019. [Online; accessed 4-June-2019].

[67] Johnson Lee. The Difference Between Systematic & Random Errors. https://sciencing.com/constant-error-12216420.html, 04 2018. [Online; accessed 16-June-2019].

[68] Cosmin Radu Crăciun. Data Management for Nudged Green Transportation. Master's thesis, UiT The Arctic University of Norway, 05 2019.

[69] Lady Limunga Jemea. Data Analysis and Nudging for Green Transportation. Master's thesis, UiT The Arctic University of Norway, 05 2019.

[70] Statens Vegvesen. About the Norwegian Public Roads Administration. https://www.vegvesen.no/en/the-npra/about-us/about-the-npra, 11 2017. [Online; accessed 4-June-2019].

[71] Statens Vegvesen. API for parkeringsregisteret. https://www.vegvesen.no/en/the-npra/about-us/about-the-npra, 03 2018. [Online; accessed 4-June-2019].

[72] Statens Vegvesen. DATEX II. https://www.vegvesen.no/om+statens+vegvesen/om+organisasjonen/apne-data/Datex, 07 2018. [Online; accessed 4-June-2019].