Jesper G. Paulsen

Making the Invisible Visible - How people can become aware of their own digital climate footprint to save the earth

Master's thesis in Computer Science
Supervisor: Svein-Olaf Hvasshovd
February 2022

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

Jesper G. Paulsen

Making the Invisible Visible - How people can become aware of their own digital climate footprint to save the earth

Master's thesis in Computer Science
Supervisor: Svein-Olaf Hvasshovd
February 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Computer Science

NTNU
Norwegian University of
Science and Technology

Abstract

The Internet is an integral part of daily life. It has revolutionized the way people interact and has laid the foundation for new technologies and information sharing. Going forward, the Internet will be essential for helping other sectors reduce their climate impact. However, the use of the Internet comes with a cost. Every piece of information sent across the Internet uses energy. Every bit of data stored in *the cloud* uses energy. Research conducted on this topic demonstrates the Internet's extensive climate footprint. However, the Internet is yet to be recognized as a mainstream climate concern.

This thesis demonstrates how individuals' Internet usage and digital climate footprint can be reduced through real-time visualization. First, a literature review is performed to understand the Internet's carbon footprint. Based on the findings from the literature review, a set of requirements are created to develop an application to visualize the users' digital climate footprint. Furthermore, this thesis presents a system that can be used to measure and visualize the digital carbon footprint of individuals.

A combination of pre/post questionnaires and an experiment is used to test the impact of the created application. The results show that the participants become more aware of their digital carbon footprint after using the application. On the other side, the results also reveal that not all the participants change their behavior when becoming aware of their digital carbon footprint. However, the percentage of the participants who change their behavior demonstrates that the application can be used to start a societal change towards a more sustainable Internet.

Sammendrag

Internett er en viktig del av hverdagen. Det har revolusjonert måten mennesker interagerer på og har vært byggeblokken for ny teknologi og informasjonsdeling. I framtiden vil internett være avgjørende for å redusere klimaavtrykket til andre sektorer. Likevel har internett en skjult kostnad. Hver eneste bit av informasjon som sendes over internett bruker energi. Hver eneste bit av data som er lagret i skyen bruker energi.

Denne oppgaven belyser hvordan internettbruk kan visualiseres i sanntid for å redusere internettets klimaavtrykk. Først gjennomføres en litteraturgjennomgang for å forstå klimaavtrykket til internettet. Basert på denne gjennomgangen lages et sett med krav for å lage en applikasjon som kan visualisere internettbrukeres digitale klimaavtrykk. Ved å se på kravene som er laget, undersøkes mulige teknologier som innfrir med kravene. Deretter presenteres et system som kan brukes for å måle og visualisere det digitale klimaavtrykket til individer.

En kombinasjon av før/etter spørreundersøkelser og et eksperiment blir brukt for å evaluere systemets innvirkning på brukerne. Resultatene viser at brukerne blir mer bevist på sitt eget digitale klimaavtrykk etter å ha brukt applikasjonen. Derimot viser resultatene også at ikke alle brukerne endret atferd selv om de ble bevisst på sitt digitale klimaavtrykk. Likevel viser prosentandelen av deltakerne som endret atferd at applikasjonen kan brukes til samfunnsendring mot et mer bærekraftig internett.

Preface

This Master's thesis was conducted during Autumn 2021 at the Norwegian University of Science and Technology (NTNU), Faculty of Information Technology and Electrical Engineering, Department of Computer Science.

First of all, I would like to thank my supervisor Svein-Olaf Hvasshovd for providing me with feedback and insight into how to execute this thesis' research.

I would also like to thank Gerry McGovern for the valuable information and research he provided during this thesis, both through interviews and emails. Without the information and details he provided in his book *World Wide Waste*, I would never have thought of the idea for this masters' thesis.

Finally, I would like to thank the participants that were part of this project. This thesis would not have been possible to execute without their feedback and participation.

Jesper G. Paulsen
Trondheim, 18th February 2022

Contents

List of Figures

List of Tables

Glossary

Digital Carbon Footprint CO_2e emissions from producing and using digital devices and infrastructure.

Green Power In this context, green power refers to energy from renewable energy sources.

Linting Linting is the process of finding code that could introduce unwanted side effects or code that do not follow a certain style guideline.

Pollution In this context, pollution refers to the emission of CO_2 equivalents.

The Cloud In this context, the Cloud refers to the servers and data centers that are accessed through the Internet.

Acronyms

API Application Programming Interface.

CAGR Compound Annual Growth Rate.

CD Continuous Deployment.

CI Continuous Integration.

CMS Content Management System.

CO_2e Carbon Dioxide Equivalent.

CRUD Create, Read, Update, and Delete.

CSS Cascading Style Sheets.

HTML HyperText Markup Language.

HTTP HyperText Transport Protocol.

ICT Information and Communications Technology.

IDE Integrated Development Environment.

IoT Internet of Things.

ISP Internet Service Providers.

JS JavaScript.

JSON JavaScript Object Notation.

LCA Life Cycle Assessment.

LCI Life Cycle Inventory Analysis.

TDD Test-driven Development.

TS TypeScript.

URL Uniform Resource Locator.

1 Introduction

1.1 Motivation

During my previous semester at NTNU, I started to think about how much we use the Internet. That is when I began to dig into the climate consequences of the Internet and found a book written by Gerry McGovern - World Wide Waste. The book made me realize that the climate impact of the Internet is vast and that we need to do something about it before it's too late.

During the 26 years I've been alive, no one has told me that I emit CO_2 equivalents every single second that I'm using the Internet. During the five years I've been studying computer science, no one has told me about how the things I deploy to the Internet pollute; everything we do on the Internet uses energy.

I feel like we today are seeing the Internet as we saw the ocean 60 years ago or so. We don't know what's happening when we upload all our images to the cloud; in the same way, we didn't realize what was happening when we threw our garbage in the ocean many years ago. But today, the negative consequences of all the trash that we have thrown into the sea are getting clearer and clearer. That's why I think it's important to focus on saving the Internet before it is too late.

Not many years ago, *the cloud* became what everyone was talking about. The cloud sounded great. We could upload everything we wanted to the Internet without thinking about where we last had seen our hard disk. We didn't have to go to the store anymore to buy larger hard disks. We didn't need to worry about losing our hard disk. Well, the cloud isn't a cloud. The cloud is just a lot of data centers stored worldwide, consuming enormous amounts of energy [81]. Every time you access the cloud, you access a data center with your files stored on their hard disks. Your data then travels from the data center, in cables, through the world back to your computer. And the worst thing is; as long as you have data stored in the cloud, the data is constantly polluting. The cloud doesn't work like your hard disk, where it is turned off when you don't need your photos. Every image, file, and video you upload to the cloud will consume energy until it is deleted. The cloud is always on.

Every time you visit a webpage, you pollute. Every time you send an email, you pollute. Every second you are watching a movie on Netflix, you are polluting. We can't stop using the Internet, and we shouldn't, but we should use it correctly so we are able to reach the climate goals (as defined by the UN of capping rise in temperatures by 1,5 degrees). We need to raise awareness around the Internet and its carbon footprint. We need to find a sustainable way to use the Internet.

1.2 Problem Description

This Master's Thesis was written during Autumn 2021, as part of the Master's Program in Computer Science at the Department of Computer and Information Science (IDI) at the Norwegian University of Science and Technology (NTNU). Together with the supervisor, Svein-Olaf Hvasshovd, a problem description was created in August 2021 to research the climate impact of the Internet. The initial problem description we created was as follows:

The goal of this project is to research, design and create an application that can aid Internet users in becoming aware of their digital carbon footprint. The goal is to make people aware of how their different online actions use energy and what actions they can perform to reduce their digital carbon footprint. Further, the project should research if people that get aware of their digital climate footprint change their behavior online to reduce their digital climate footprint.

1.3 Reader's Guide

The thesis is structured as follows: First, the research goals and questions are presented. Then some background information is presented. In Chapter 3, the research method is given. Next, Chapter 5 presents the carbon footprint of the Internet. Following, Chapter 6 highlights the technical research conducted to understand how the carbon footprint of individuals can be calculated. Then, Chapter 7 explores how other applications are used to visualize behavior and the carbon footprint of their services. The requirements of the application are then outlined in Chapter 8. Chapter 9 presents the planned architecture of the system, as well as the potential technology for each of the components of the Architecture. Next, Chapter 10 describes the implementation of the application, and Chapter 11 validates the requirements of the application. Following, Chapter 12, outlines the execution of the data generation methods. The results of the data generation methods are then presented in Chapter 13. In Chapter 14, the findings from both data generation methods are discussed, as well as the limitation of this research and possibilities for future work. Finally, Chapter 15 provides a conclusion to this thesis.

2 Research Goal and Research Questions

This chapter will introduce the research goal and research questions of this thesis. The research goal and research questions have been created based on the Goal Question Metric approach (GQM) introduced by Basili [25]. The GQM approach was initially designed for several NASA Goddard Space Flight Center projects [25]. The research method characterizes a top-down approach to define metrics for measuring and giving feedback on applications. The GQM approach has three levels: First, an overall goal is defined. Second, a set of questions are defined to characterize the achievement of the goal. At the third level, metrics are defined to answer the question.

2.1 Research Goal

The overall research goal of this thesis is defined as:

Characterize the effect visualizing an individual's digital carbon footprint has on the digital carbon footprint of the individual and the total carbon footprint of the Internet in the future.

The goal is to look at the Internet's carbon footprint and create an application to make individuals aware of their digital carbon. To answer the research goal, a proper understanding of the Internet and its climate footprint is necessary. It will likewise be crucial to comprehend the technical aspect of such an application and how the data can be visualized to the end-user to aid them in understanding their digital carbon footprint.

2.2 Research Questions

In order to achieve the research goal, a set of research questions have been created.

2.2.1 RQ 1: What is the carbon footprint of the Internet today and what are the trajectories for the future?

In order to create an application that will make people aware of their digital carbon footprint, it is essential to understand the carbon footprint of the Internet and the trajectories for the future.

2.2.2 RQ 2: How can an application used to estimate the digital carbon footprint of a user be created without significantly increasing the digital carbon footprint of a user?

A crucial part of this thesis is to understand the data needed to estimate the carbon footprint of individuals. Second, it is essential to research how to develop an application that is able to collect the data required to estimate the carbon footprint of individuals. Third, it is vital to explore if it is possible to estimate and give feedback on the digital carbon footprint of an individual without significantly increasing the digital carbon footprint of the individual.

2.2.3 RQ 3: How can people become more aware of their digital carbon footprint?

It will be prominent to explore how the carbon footprint of individuals can be visualized to aid them in becoming aware of their digital carbon footprint.

2.2.4 RQ 4: Will people that become more aware of their digital carbon footprint change their behavior on the Internet?

As mentioned by Wamsler et al, behavioral change is necessary to reduce the impact of climate change [145]. Therefore, it is necessary to research if people who get access to see their digital carbon footprint will reduce their digital climate footprint.

2.3 Summary

This chapter has presented the Goal Question Metric approach used in this thesis. Four research questions have been created to reach the goal of this thesis: *Characterize the effect visualizing an individual's digital carbon footprint has on the digital carbon footprint of the individual and the total carbon footprint of the Internet in the future.*

3 Background

This chapter aims to provide the reader with the background information for this thesis. First, climate change will be addressed. Second, the technical details of the Internet and the Web will be discussed. Next, the thesis uncovers common details of today's digital habits. Finally, an essential paradox, namely the *Jevons' Paradox*, will be explained.

3.1 Climate Change and the ICT Sector

In 2021, IPCC, the Intergovernmental Panel on Climate Change, released its sixth assessment report. This report concludes that there is no doubt that the climate changes we are seeing are caused by human activity [13]. Climate changes pose a significant threat to humanity [96], and the consumption of energy is one of the main drivers behind climate change [2].

Technology has changed the way humans live. An example of this is that the Internet has made it possible for anyone with Internet access to stream almost any video anywhere in the world. At the same time, research shows that the ICT sector's carbon footprint (2.1-3.9% of the World's total emissions) [56] is roughly the same as the aviation industry (3.5% of the World's total emissions) [113]. If the ICT sector is to meet the 1.5 °C trajectory needed to reduce the impact of the climate changes, the emissions from the ICT sector need to be 51% lower in 2030, compared to 2015 [139].

Research shows that only technological inventions and public governance alone cannot stop climate change but also requires behavioral change by individuals [145]. To that end, individuals' behavioral change to reduce their carbon footprint is being observed to a larger extent than earlier. People are increasingly choosing train instead of planes for transportation, and vegetarian diet instead of meat [67]. However, individuals are at the same time consuming and producing more and more data [38].

3.2 The Internet

3.2.1 Architecture

The Internet is a complex system consisting of billions of different devices and components. A simplified version of the Internet is shown in Figure 3.1.

User Devices Includes all types of devices a end-user uses to browse the Internet.

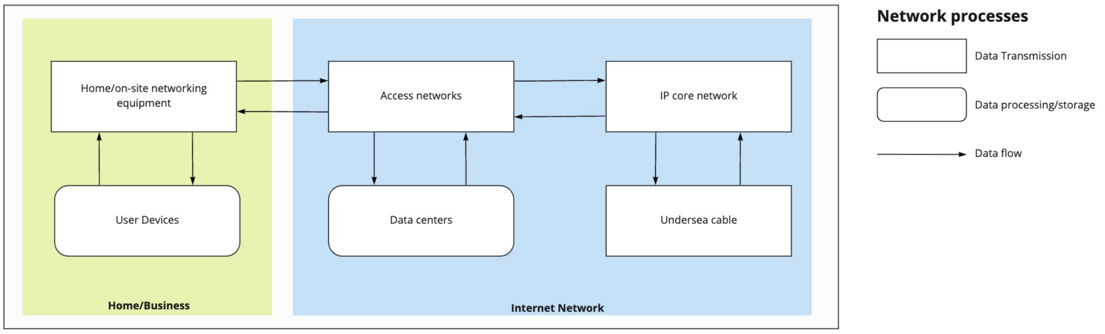

Figure 3.1: A simplified figure of the Internet's Architecture, based on Aslan et. al Figure of the Internet [12].

Home/On-site Networking Equipment includes components that lets the user connect to the Internet. This is often a home router or a base station providing 4g and 5g. This is also called Customer Premise Equipment.

Access Networks is the equipment that connects the customers to the Internet Service Providers (ISP). An Internet Service Provider is a company that offers private persons and companies Internet access. They send the data from the customer in cables to the IP Core network and back.

Data Centers are buildings and infrastructure that house large servers. These servers can be used to process or store data. Data centers use water and air-condition to cool down their components. The data centers are placed worldwide, but the majority are located in the USA [21]. Most websites and cloud services are accessed through a data center.

IP Core Networks are the Internet components that make up the regional, global and national Internet networks. These are transmission cables, links, routers, and switches. They move data in cables between the Access Networks and Undersea Cables.

Undersea Cables are the cables sending data between continents and countries. A map of the undersea cables can be found at the Submarine Cable map[1]

3.3 The Web

It is essential to understand the difference between the Internet and the Web when talking about the Web. The Web is the collection of data that is accessible through the Internet. There are other forms for Internet traffic, but the Web is the most dominating [11]. The Internet is the infrastructure that makes it possible to access the Web. A more

[1] `https://www.submarinecablemap.com/`.

straightforward analogy is to think of a library as the Internet, and the books in the library as websites, making up the Web. The library provides all the infrastructure that is required to store the books, and it provides the infrastructure required to access the books. In the same way, does the Internet provide access to the Web. The Web consists of the information created by the authors (Developers). The information is used by the end-users (Consumers). In 1994 there where 3000 websites world wide [129], whereas in 2021 it was estimated to exist more than 1.88 billion websites [129].

3.3.1 Network Call

Network calls are required to exchange information between an internet browser and a server. A network call, also known as a network request, consists of two parts: A HyperText Transport Protocol (HTTP) request, and a HTTP response [100].

A HTTP request goes from the consumer's device (client) to a host on a server [70]. It consists of a *Request Line*, *HTTP Headers* and an optional *Body*.

Request Line The request line contains the HTTP method used, the path to the requested resource, and the HTTP version. The HTTP method determines what kind of action the server should perform on the request [102], and the path determines what resource the client requests. Finally, the HTTP version tells the server what HTTP version the client has used [98].

HTTP Headers The HTTP Headers contain information that the server needs to process and respond to the client's request [101]. Each header has a name and a value.

Body The body is optional but can contain information the server needs to process a request.

A HTTP response goes from the server to a client, in response to a HTTP request [71]. It consists of a *Status line*, *HTTP Headers* and an optional *Body*.

Status Line The status line contains the HTTP version used by the server, the status code, and a status text. The status code indicates the result of the request. Different status codes have different meanings, and some status codes are 404 and 200 [103]. The status text is a readable explanation of the status code, "Not Found" for the status code 404 and "OK" for the status code 200.

HTTP Headers The HTTP Headers are similar to the ones that go from the server to the client, containing information the client might need to process the request. This can be information about the server, or information used by the client to show information to the user.

Body The message body in the HTTP response is also optional, but is used for most responses [71]. This body can contain a resource the client requested or an indication of why a request failed.

3.3.2 Page Weight

Page Weight is the size, in bytes, of a web page. A web page consists of several elements, and all of them affect the weight. The larger the element is, the more data will be transferred from the webserver to the end-user, and the more energy is required to process and render the element on the end consumer's device.

Assets Assets can have a significant impact on the page weight. Assets are images and videos on the web page. Videos are one of the Web's largest asset type and require much energy when being transferred from a server to an end-user device. Images will also affect the page weight, and images were in 2021 on average 70% of the total page weight of a web page [10].

JavaScript (JS) JavaScript is the programming language that is used when developing modern websites [73]. It is the bare-bone of interactive web pages and lets the developer dynamically show elements to the consumer. In 2021 JavaScript was responsible for, on average, 20% of the page weight. On computers, 648 KB of JavaScript was loaded on average, and on mobiles 589 KB [10]. In addition to using energy when transferred and stored on the server, JavaScript can cause high energy usage on the end-user device.

Fonts Fonts will also impact the page weight. In 2021 fonts was responsible for, on average, 5.5 % of the page weight [10]. The font size is determined by several factors, including the number of characters and instructions on displaying the font to the end-user.

Cascading Style Sheets (CSS) CSS is used to style web pages. Colors, spacing, and size are styling attributes that are controlled by CSS [142]. In 2021, CSS was responsible for, on average, 4% of the page weight of a web page [10].

HyperText Markup Language (HTML) HTML is the most basic building block for web pages [97]. HTML structures websites and defines where text and assets should be shown. It also defines which elements should get CSS and JavaScript applied to them. In 2021, HTML was responsible for, on average, 1.8% of the page weight of a web page.

Network Requests Every time a web page is visited, network calls are sent to servers from the consumer's device. It might be enough for a simpler web page with only one network call, but the number of network calls can be several hundred for more complex sites. The number of network calls will also depend on how the webpage is bundled. With modern bundlers like WebPack, it is possible to split the JS code into smaller chunks

[148]. This has the advantage of faster loading of the required modules but will require several network requests.

Tracking When a consumer visits web pages, the consumer is tracked by "first parties" and "third parties" [44]. The first parties are often tracking tools for monitoring the performance and stability of the site. In contrast, third parties tracking are hidden ad-networks embedded in the web page [44, 57]. Despite laws that enforce online privacy, tracking is still a common practice on websites [125, 115]. Tracking makes it possible for third parties to build up a consumer's web history and a more complex picture of the consumer. Tracking to both first and third parties services will increase network requests and page weight. All the tracking data will also contribute to more data being stored.

3.4 Internet Habits

3.4.1 Amount of Data Sent Across the Internet

By 2022, it is expected that more than 4.8 zettabytes [2] are transmitted across the Internet annually [18]. Due to the Covid-19 pandemic, the Internet traffic is estimated to be even higher [49, 24, 91]. Feldman et al. and Bottger et al. found that the Internet traffic increased with 15-20% in a couple of weeks during the pandemic, and Bottger et al. reported that the greatest increase was seen when the largest countries in the world implemented lockdowns. Normally this increase was seen during a couple of months [49].

Both papers saw a tremendous increase in data traffic from video services after the beginning of the pandemic. Feldman et al. conclude that the data traffic from web conference tools increased 200% during the pandemic. Bottger et al. observed an increase of 250%-350% in data traffic from live streaming.

3.4.2 Time Spent Online

Similar to the increase in data traffic over the Internet, there has also been a steady increase in time spent on the Internet [38, 136, 111]. The time spent online per country varies, and Statistics Norway (SSB) reports that the average time spent online per day in 2020 by Norwegian citizens to be 3 hours and 22 minutes [136]. Similarly, Ofcom, the regulator for communications services in the UK, reports that time spent online per day for UK citizens in 2020 was 3 hours and 37 minutes [111]. When looking at the average of all countries, DataReportal, reports that the average time spent online per day in 2021 was 6 hours and 58 minutes [38].

3.4.3 Number of People and Devices Online

The number of people online is steadily increasing. In 2021 it was estimated to be between 4.7 and 4.8 billion people connected to the Internet [30, 38]. By 2023 it is expected to

[2] 1 zettabyte is equal to one trillion gigabytes, or 10^{21}. 4.8 zettabytes in one year equals all movies ever made being sent across the Internet every 53 seconds for a year. [18]

be 5.3 billion people connected [30]. This makes up a compound annual growth rate (CAGR) of 6%. The number of devices connected to the Internet is estimated to be just below 25 billion in 2021, and in 2023 it is expected to be just below 30 billion [30]. This equals a CAGR of 10%.

3.4.4 Internet Speed

The average internet speed is increasing. In 2021, Cisco predicted that the average fixed broadband speed was 77.4 Mbps, and the average mobile speed was 29.4 Mbps [30]. However, DataReportal reports the average fixed broadband speed in 2021 to be 106.61 Mbps and the mobile speed to be 55.34 Mbps [38]. In similarity, Speedtest.net[3] reports that the average fixed broadband speed in 2021 was 107.50 Mbps, and 55.07 Mbps for average mobile speed [112].

3.5 Jevons' Paradox

When researching Internet habits and the future of the Internet and its climate footprint, it is relevant to mention Jevons' Paradox. William Stanley Jevons presented Jevons' Paradox in "The Coal Question: an Inquiry Concerning the Progress of the Nation, and the Probable Exhaustion of our Coal-mines.", published in 1866 [75]. In his paper, Jevons described how the introduction of coal led to even higher consumption of coal, iron, and other resources. In a paper from 2005, Blake Alcott explained how twentieth-century economic growth theory, in similarity to the phenomenon Jevons described, saw the growth in consumption and production as a result of technological shift [4]. The study also reported that "In contrast, some ecologically-oriented economists and practically all governments, green political parties and NGOs believe that efficiency gains lower consumption and negative environmental impact." [4]. Furthermore, a paper called "Jevons' paradox revisited: Implications for climate change" was released in 2021. The authors of this paper conclude that improved fuel efficiency always raises the total stock of carbon emitted [128]. It is possible that Jevons' Paradox also holds for the ICT sector and that a more efficient Internet will lead to more data consumption.

3.6 Summary

This chapter has presented the connection between the Internet, the Web, and Climate change. It has described how a web page's elements affect its page weight, affecting its climate footprint. Further, it has looked at the Internet habits of users and the forecasted usage. It has also adressed the Jevons' Paradox, which may serve as an indication that we will continue to consume even more data in the future.

[3] `https://www.speedtest.net/`

4 Research Method

This chapter will outline the Research Approach of this thesis. An overview of the chosen research steps for this project is illustrated in Figure 4.1, based on the research model presented by Oates [109]. The green boxes indicate the chosen research steps conducted during the execution of this research. To the author's knowledge, researching how users can become aware of their digital carbon footprint has not been conducted before. As the area of this thesis is new, a mixed-method approach was chosen.

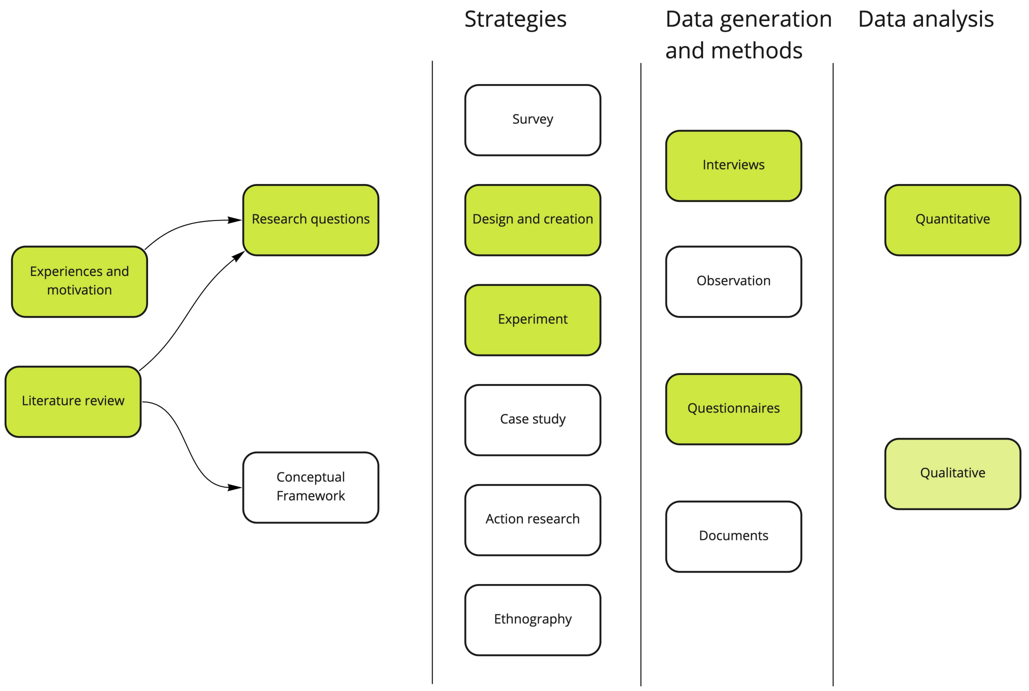

Figure 4.1: The Research Model based on Oates research model [109]. Chosen steps for this thesis are highlighted in green.

4.1 Preparation

This thesis's preparation, research, and execution were performed during Autumn 2021. Prior to this thesis, the book World Wide Waste[1] was read, which provided context for the starting point of the thesis.

4.2 Literature Review

The first phase of the research conducted consisted of a literature review. The literature review aimed to understand how Internet usage affects the Internet's climate footprint. Google Scholar[2] was used to identify existing literature and applications within this research field.

In combination with a literature search, snowballing was used. Snowballing methods has shown to be as successful as database searches [150]. When performing the initial article search, I used keywords such as *Climate change*, *Internet's climate footprint*, *ICT energy usage*, *Digital behavior* and *Digital carbon footprint induviduals*.

4.3 Experts and Interview

As the research field is relatively new, an interview with Gerry McGovern was conducted. The researchers Anders Andrae and Jens Malmodin were also emailed to get a better overview of the Internet's carbon footprint. They have both written multiple papers [7, 5, 6, 86, 85, 84] in the research field and provided valuable information about the energy usage of the Internet.

4.4 Technical Research

After getting an overview of the Internet's energy usage and how individuals' digital climate footprint can be computed, technical research was performed. This part focused on the potential solutions that could be utilized when developing an application to capture and estimate the digital carbon footprint of individuals. The research performed during this part was crucial for the rest of the thesis.

4.5 Creation

In this phase, the proposed solution was designed and implemented. First, relevant applications were examined to gather some inspiration. Based on the insights from the literature review, the technical research, the relevant applications, and the interview, a set of requirements and sketches were produced. Then, the architecture was planned. A compromise was made between the carbon footprint of the architecture and the time

[1] https://gerrymcgovern.com/world-wide-waste/
[2] https://scholar.google.com/

required to make the application as environmentally friendly as possible. Next, the initial version of the application was created. This version was used for user testing to identify aspects that needed improvements. Eventually, the final implementation of the application was developed.

4.6 Sampling of Participants

The chosen sampling method for this project was *Self-selection sampling.* Self-selection sampling is a sampling technique where the researcher advertises the research project, and the participants sign up themselves. This is a method used when the researcher does not know how to get in touch with people that might want to participate but might lead to a sampling pool where most of the participants have a strong feeling on the subject [109]. I posted a post on LinkedIn with information about the project to advertise the project. Gerry McGovern also shared the project on his Twitter account.

4.7 Data Generation Methods

The thesis uses two data generation methods; pre/post Questionnaires and Observation. After the participants signed up, they received a questionnaire. This questionnaire was used to explore the participants' internet habits and their digital carbon footprint. Next, an experiment was executed over a time period of two weeks. During the first week, the application stored the participants' estimated carbon footprint, but the participants did not have access to the estimations. The second week, the participants were able to explore their estimated digital carbon footprint in real-time. After the experiment, a final questionnaire was answered. This questionnaire aimed to understand the participants' thoughts about the application and their digital carbon footprint.

4.8 Data Analysis and Discussion

The data collected through the application and the results from the two questionnaires were analyzed. The results from the data analysis, the literature review, and the application developed were discussed. Finally, the limitations of the thesis were examined, and focus points for future work was created.

5 The Carbon Footprint of the Internet

This chapter will introduce the carbon footprint of the Internet. It will highlight the research conducted in the field and summarize some of the studies published between 2019 and 2021. Then, the chapter will pinpoint some of the challenges that exist when estimating the ICT sector's carbon footprint. Based on the findings presented in this chapter, a conclusion regarding the Internet and its climate footprint will be presented.

5.1 Research Estimating the Carbon Footprint of the Internet

Since 2000, multiple researchers have been trying to estimate the Internet's electricity usage and carbon footprint, and in 2017 a paper by Aslan et al. was published [12]. This study was based on multiple studies from 2004 to 2015 [79, 134, 17, 147, 35, 149, 82, 86, 36, 127, 126, 68, 85]. The study found that since 2000 the estimations for energy per byte, kWh/GB, had varied with more than five orders of magnitude; from 136 kWh/GB estimated by Koomey et al. in 2004 [79], to 0.004 kWh/GB estimated by Baliga et al. in 2008 [17].

The studies Aslan et al. examined had various system boundaries. Therefore, Aslan et al. only focused on the transmission network system, which consists of the IP-Core Network and the Access Network, as presented in Section 3. Using a mixture of regression and extrapolation on the six studies that met their criteria, they found that the electricity intensity for the transmission network was halved every second year. They also concluded that their research could help identify the electricity intensity for the years between 2000 and 2015, but "More research is required to update estimates for current and future years" [12]. For 2015, they calculated an electricity intensity of 0.06 kWh/GB.

Since 2017, multiple research groups have looked at the carbon footprint of the Internet [84, 6, 124, 80, 72, 23, 87, 139, 56, 135, 69, 95, 88, 41, 59, 3]. As with earlier reports, the estimations still vary. I will now present some of the articles and their findings.

The Shift Project and The International Energy Agency In 2019, the Shift Project[1], a so-called "think-tank" from France, released a report "Climate crisis: The unsustainable use of online video – A practical case study for digital sobriety" [41]. The paper reported that "the emissions generated by watching 30 minutes of Netflix [1.6 kg

[1] `https://theshiftproject.org/`

of CO2] is the same as driving almost 4 miles" [41]. This quote received much attention in media, even in 2021 [37, 89, 83, 108, 42, 16, 1].

The International Energy Agency[2], IEA, released in 2020 an article called "The carbon footprint of streaming video: fact-checking the headlines" [3]. The article's goal was to point out that the number that the Shift Project and different media were using was too high. The IEA estimated the data transmission energy to be 0.019 kWh/GB [3].

In response to the IEA report, The Shift Project acknowledged that they had been using bits instead of bytes in their estimations, which resulted in wrong calculations. They later published their corrected calculations, where they estimated the data transmission energy to be between 0.15 and 0.88 kWh/GB [59, 41].

Malmodin 2020 [84] Malmodin has released multiple papers on ICT and its carbon footprint. As a researcher at Ericsson[3], he has access to multiple data sources. Furthermore, he released two papers [86, 85] whom both were part of Aslan et al. paper from 2015. His paper "The power consumption of mobile and fixed network data services - The case of streaming video and downloading large files" released in 2020, focuses on "Numerous claims in media about the electricity consumption of video streaming and data downloading over mobile and fixed networks." [84]. The study proposed a new way to estimate the Internet's energy usage; instead of looking at how much energy the Internet uses per data (kWh/GB), Malmodin suggested that it will be more accurate to look at energy per time. This is a more modern approach, called the Power Model, where the baseload of the network is allocated per subscriber/user, and the different network component is allocated based on the data used.

Through email, he elaborated on the findings presented in the paper. To argue why he found it more appropriate to use energy per time as measurement units, he used a 4g base station as an example. The baseload for a 4g base station is around 200W for 200 connected devices. A marginal increase in energy usage is noticed per extra device connected to the base station. Connections over the top will only add around 1-2 W/Mbps. "The power consumption does not start when we use it. It is there all the time." [84]. Malmodin estimated that the carbon footprint, relative to the global carbon footprint, of the ICT sector peaked in 2010 and has been declining since.

Andrae 2020 [6] Andrae is a researcher at Huawei[4], and he has published multiple papers in this field [7, 5, 6]. His papers focus on forecasting the ICT sector's carbon footprint. In his paper "New perspectives on internet electricity use in 2030," he estimated the global electric power use of the ICT sector. He compared his findings to the paper released by Andrae and Edler in 2015 [7]. In general, he found that his estimations were lower for both 2020 and 2030 than the findings in the paper from 2015.

When discussing the issue with Andrae (by email), he pointed out that he estimated 0.09 kWh/GB for regular internet browsing in 2021. He also argued that he still sees

[2] https://www.iea.org/
[3] https://www.ericsson.com/en
[4] https://www.huawei.com/en/

kWh/GB for the whole Internet divided by the Global Data Center IP traffic as the best approximation for the Internet's electricity intensity. [6].

Obringer et al. 2021 [110] Obringer et al. released in 2021 a paper that focused on how the fourth industrial revolution and the COVID-19 crisis accelerated the transition to an "unregulated and environmentally unaudited digital world" [110]. In contrast to most research conducted in this field, they also included calculations for the ICT sector's water and land footprint. They estimated that the water footprint would be enough to fill 317,200 Olympic size swimming pools and that the land footprint was the size of Los Angeles [110]. By the end of 2021, they estimated that the global carbon footprint from the ICT sector could grow to 34.3 million tons of CO_2 emissions.

They estimated a median carbon footprint of 32.13 g CO_2e/GB. Dividing this number with the world carbon intensity of 475 g CO_2e/kWh [43] yields 0.068 kWh/GB. Their findings conclude that Internet users can reduce their footprint by 86% if streaming in standard definition instead of high definition. Further, they found that Internet users can reduce their footprint by 96% by turning off their cameras in digital meetings.

Ruiz et al. 2022 [124] Ruiz et al. published in 2022 a paper that estimated the carbon footprint of wireless ICT networks (4g LTE) for six demographic areas in Peru which did not have such services before the installation. Embodied emissions from global transmission cables and data centers were left outside of the system boundaries for their study. They estimated 1.35 - 1.73 kg CO2e /GB. Their study reported that most of the emissions come from the end-user (68-86%), where the embodied emissions account for the biggest part [124].

International Telecommunication Union 2020 [139] The International Telecommunication Union (ITU)[5] is the United Nations specialized agency in the field of ICT. In 2020 they released a report called "Greenhouse gas emissions trajectories for the information and communication technology sector compatible with the UNFCCC Paris Agreement." The goal of the report was to look at how the ICT sector needs to evolve to stay in line with the 1.5 °C scenario mentioned in the IPCC report. They used data from 2015 as the baseline year to model how the ICT sector will develop with and without reduction compared to the 1.5 °C trajectory. To meet the target in the Paris Agreement, they estimated that the ICT sector needs to reduce its emission by 51% by 2030 compared to 2015.

Carbon Trust 2021 [135] Carbon Trust[6] is a company specialized in assisting businesses, governments, and organizations in decarbonization [135]. In 2021 they published a paper called "Carbon impact of video streaming," estimating the carbon footprint of streaming video for one hour in Europe. The white paper was made available

[5] https://www.itu.int/
[6] `https://www.carbontrust.com/`

by funding from Netflix. They estimated that one hour of streaming would emit 55 gCO2e/hour. They also found that the end-user device is responsible for 28 gCO2e/hour (=51%). With the introduction of the "Right to Repair" program that the European Parliament voted for in 2021, they believed that the end-user devices' carbon footprint could be reduced. "... the ICT and E&M sectors face various challenges to decarbonize. However, through the power sector and purchases of renewable electricity, ICT is well-positioned to keep pace with future targets." [135].

Marks et al. 2021 [87] Marks et al. published in 2021 a report called "Tackling the Carbon Footprint of Streaming Media." In their research, they reviewed 170 articles, compared 22 calculations, and surveyed seven existing calculations of the energy consumption of ICT. Their triangulation established that most ICT researchers arrive at estimates similar to The Shift Project's. They concluded that video streaming is responsible for over 1% of greenhouse gas emissions worldwide and that an absolute decrease in energy consumption is necessary [87].

Freitag et al. 2021 [56] Similar to Marks et al. [87], and Aslan et al. [12], Freitag et al. estimated the climate impact of the ICT sector based on earlier papers. They reported that all the calculations they reviewed underestimated the carbon footprint by as much as 25% by "failing to account for all of ICT's supply chains and full lifecycle." [56]. By adding the entire lifecycle to the calculations, they estimated that the ICT sector is accountable for as much as 2.1-3.9% of the global GHG emissions. They concluded that major efforts from both the industry and politicians are essential to lower the emissions from the ICT sector.

5.2 Variations in the research

In 2021, Billstein et al. released a paper aiming to pinpoint the reasons for the significant variations in the estimations of the Internet's carbon footprint [23]. They identified eight possible challenges:

- **System boundaries**: They found that there is no standardization of system boundaries. They discovered that the most common system boundary included the Access Networks, Home/on-site networking equipment, and the IP-core network, excluding the user devices and datacenters. They argued that the system should be modeled as three modules: User devices, network equipment, and data centers [23].
- **Data Collection Methods**: Collecting the data necessary to estimate the carbon footprint of the Internet is intricate, and it is difficult to avoid using generic data. They concluded that the best way to handle this problem is to use a combination of a Top-Down and a Bottom-Up approach [23].
- **Measurement Units for Electricity Intensity**: Researchers are using various approaches when estimating the average electricity intensity, such as energy per

data, energy per time, and energy per subscription. Their research deduced that the access network, customer premise equipment, and end-user devices are best modeled using energy per time. Further, the IP core network should be modeled using energy per data [23].

- **Transparency and Data Availability**: Data that is required to estimate the carbon footprint of the Internet is not publicly known, which implies that the researchers have to make estimations when calculating the carbon footprint of the Internet. They locate that more research needs to be performed in this area, but that the data should be shared by free will, and that the increasing interest in the carbon footprint of the Internet is making the stakeholders of the Internet more aware of the positive outcomes of sharing that data [23].
- **Age of Data**: Increased improvements in the efficiency of Internet components affect the energy used by the Internet. This makes it challenging to understand which numbers are used by researchers. Older estimations also need to be continuously improved to account for the efficiency gains in the ICT sector [23].
- **Allocation Procedures**: Allocating the Internet's energy consumption is a complicated procedure due to the multifunctionality of the Internet [23]. The allocation must be made between all the components of the Internet, both the ones processing Internet traffic and the ones that do not process Internet traffic. As pointed out by Malmodin, the network components are using energy even though they are not handling Internet traffic. More research needs to be conducted in the field to determine how to allocate the energy usage [23].
- **Assumptions during LCI Phase**: Life Cycle Inventory Analysis (LCI) is the procedure of identifying and quantifying all resources used to produce a given product [106], such as a Router. Variations in the energy required to produce components and the electricity mix used during production and operation of the product can have a significant impact when estimating the overall energy usage of the Internet [23].
- **Limited Coverage of Impact Categories**: Most of the research on the Internet's climate impact has concentrated on the carbon footprint of the Internet, without examining the other environmental impacts. Billstein et al. concluded that more research needs to be conducted to understand how other environmental impacts, such as the water- and land footprint, can be included in the research.

The challenges identified by Billstein et al. can aid immensely in understanding the variation in the estimations of the Internet's carbon footprint.

5.3 Conclusion

Based on the results presented in the previous Section and the findings of Aslan et al., Marks et al., and Freitag et al., it is clear that the estimations of the Internet's carbon footprint vary extensively. Obringer et al.'s eight challenges provide great help in understanding why the estimations vary.

During the research, it has become evident that there is no clear answer to how much energy the Internet uses, and it is hard to estimate its climate footprint. However, it is also apparent that the ICT sector needs to reduce its energy usage, to be able to meet the 1.5°C trajectory proposed by the Paris Agreement. In order to meet the goals, the ICT sector needs to make sure to overcome the challenges proposed by Billstein et al.

6 Technical Research

In this chapter, the technical research of the thesis will be presented. Research is conducted to discover possible solutions that can be used when estimating the digital carbon footprint of Internet users. Then, the possibilities and limitations of the potential solutions are examined. Finally, the selected approach is presented.

6.1 The Challenge

As presented in Chapter 5, the energy usage and carbon footprint from Internet usage is complex to measure and estimate. Multiple aspects affect the carbon footprint, and there is not a single mathematical formula that can be used to calculate the exact carbon footprint of the Internet. As presented in Chapter 5, different metrics are used when measuring the emissions from Internet usage. For this thesis, the two applicable methods are the conventional method (kWh/GB) and the power model approach [84, 135]. The power model approach does not include data centers' energy usage and requires specific calculations based on the data flow through the different components. On the other side is the conventional method based solely on the energy usage per amount of data. Thus, the conventional method is simpler to implement.

For this thesis, the decision was made to calculate the emissions based on the conventional method. First, the amount of data is converted to energy using a factor of 0.09 kWh/GB. This factor was chosen based on the input from Andrae and the research conducted by Obringer et al. [110]. Then the estimated CO_2 equivalents are calculated based on the carbon intensity of the destination country of each network request.

6.2 Technical Requirements

When creating an application to estimate the carbon footprint of individuals' Internet usage, several approaches can be taken. Before deciding on the framework to use, functional and non-functional requirements are formed. The functional requirements are made to understand the functionality required by the system. The non-functional requirements are created to get a better overview of the attributes and constraints on the system.

The functional requirements are presented in Table 6.1. The non-functional requirements are presented in Table 6.2. FR1, FR2, and NFR 1 were prioritized among the Requirements during the technical research.

ID	Functional Requirement	Priority
FR 1	The system should detect amount of data being sent	High
FR 2	The system should detect the location the data is being sent to	High

Table 6.1: Functional Requirements created for the Technical Research

ID	Non-Functional Requirement	Priority
NFR 1	Implementation time less than 3 months	High
NFR 2	The system should be secure	High
NFR 3	The system should be runable on Windows, Mac OS and Linux	High

Table 6.2: Non-Functional Requirements created for the Technical Research

6.3 Possible Solutions

Three possible approaches are researched. The different approaches are evaluated based on the requirements created for the system.

6.3.1 Desktop Application

One of the possible approaches is to create a desktop application. A desktop application can monitor all network traffic that is sent from the computer to the Internet. An example of a desktop application that monitor network traffic is Whireshark[1]. The interface of Whireshark is shown in Figure 6.1.

Pros An advantage of implementing the system using a desktop application is that it will monitor all network requests, both from web browsers and native programs. Thus, this will give the user a comprehensive overview of the programs using the Internet and the estimated CO_2 emissions from such programs. This will also give the exact size, and IP address of the network calls being sent from the computer. Hence, FR1 and FR2 will be fulfilled.

Cons Even though a desktop application will be able to monitor all the network traffic being sent from a computer, the other applications will be black boxes, out of the control of the application. The application will, for instance, not be able to monitor what web pages a user visits through a web browser. Hence, the system's ability to give users feedback on how their behavior affects their digital carbon footprint will be limited.

For the application to comply with NFR 3, the system has to be executable on all operative systems. As NFR 1 limits the time scope of the application of the system, I researched programming languages I already had experience with that supported cross-platform execution. These were .NET[2] and Flutter[3]. Both these frameworks support

[1] `https://www.wireshark.org/`
[2] `https://dotnet.microsoft.com/en-us/`
[3] `https://flutter.dev/`

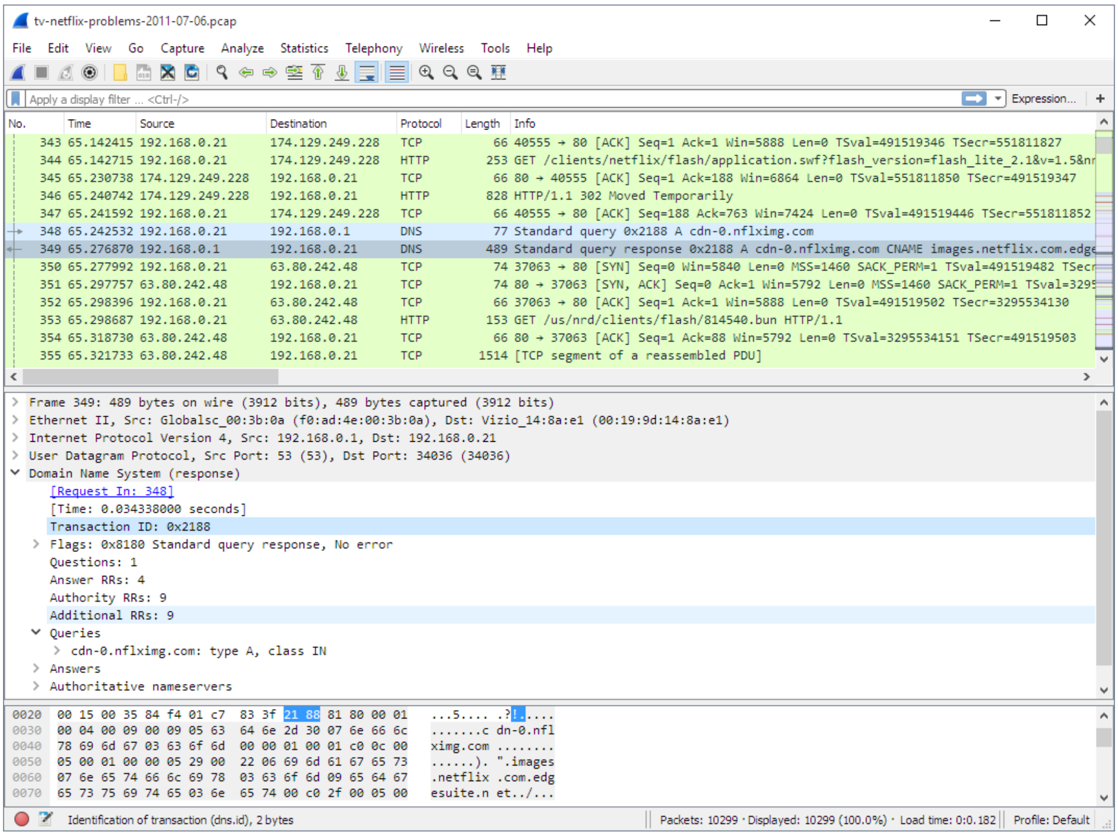

Figure 6.1: The interface of Whireshark, used to monitor network traffic. Fetched from `https://www.wireshark.org/docs/wsug_html_chunked/ChUseMainWindowSection.html` (Accessed 11/12/2021)

monitoring network calls, but none of them provides the APIs required to monitor network calls natively. Thus, implementing the application using one of the languages requires writing specific code for each operative system. This is assumed to take more time than the three months specified in NFR 1.

Desktop Browser Initially, before doing any research, the plan was to create a desktop browser. The chosen framework for implementing a browser was Chromium[4], which is the open-source project behind the Google Chrome Browser [29]. After doing some research, and in order to comply with NFR 1 and NFR 2, the approach was shelved as it would require a an amount of work not viable under the time constrains of the assignment [31].

[4]`https://www.chromium.org/`

6.3.2 Mobile Application

A mobile application can also be chosen as the selected approach. However, unlike a computer, the operative systems on mobiles do not allow applications to monitor network traffic from other applications. Thus, to capture network calls on a mobile, the only applicable method will be to create a web browser for iOS and Android. Luckily, there exists a plugin to Flutter that exposes the required APIs to create a web browser [117, 116]. An example of a browser created with this plugin is shown in Figure 6.2.

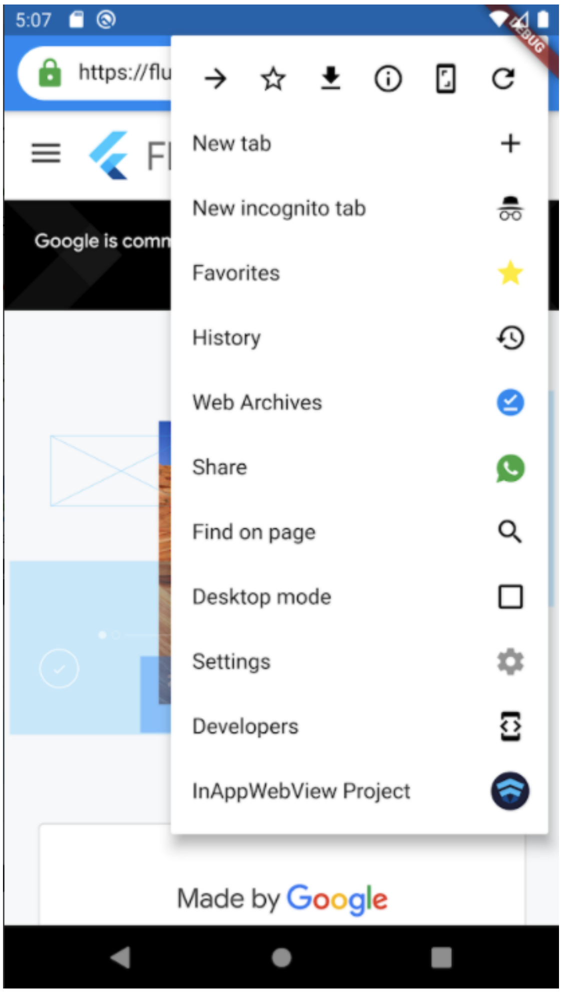

Figure 6.2: Flutter Browser: An Internet browser created using the Flutter plugin *InAppWebView*. Its design is greatly inspired by Google Chrome for mobiles. Fetched from `https://play.google.com/store/apps/details?id=com.pichillilorenzo.flutter_browser` (Accessed 11/12/2021)

Pros In 2021, 63% of page visits was through mobile, and 32% of page visits was through desktop browsers [130]. Creating a mobile browser is therefore believed to capture a significant share of all the web traffic.

Cons As a mobile web browser will not capture traffic from other applications running on the mobile, the amount of data traffic captured will be limited. Creating a browser using the plugin with high usability will also take time and can require more time than the three months specified in NFR 1.

6.3.3 Web Browser Extension

The third option researched was a web browser extension. A web browser extension only requires a host browser to run on and gets access to the API exposed by both the host browser and a normal web page. Its created using the same technology as a webpage - HTML, CSS, Assets and JavaScript [34]. Google Chrome, Mozilla Firefox, Safari, and Microsoft Edge share most of the same extension API [8, 62]. An example of a Web Browser Extension is shown in Figure 6.3.

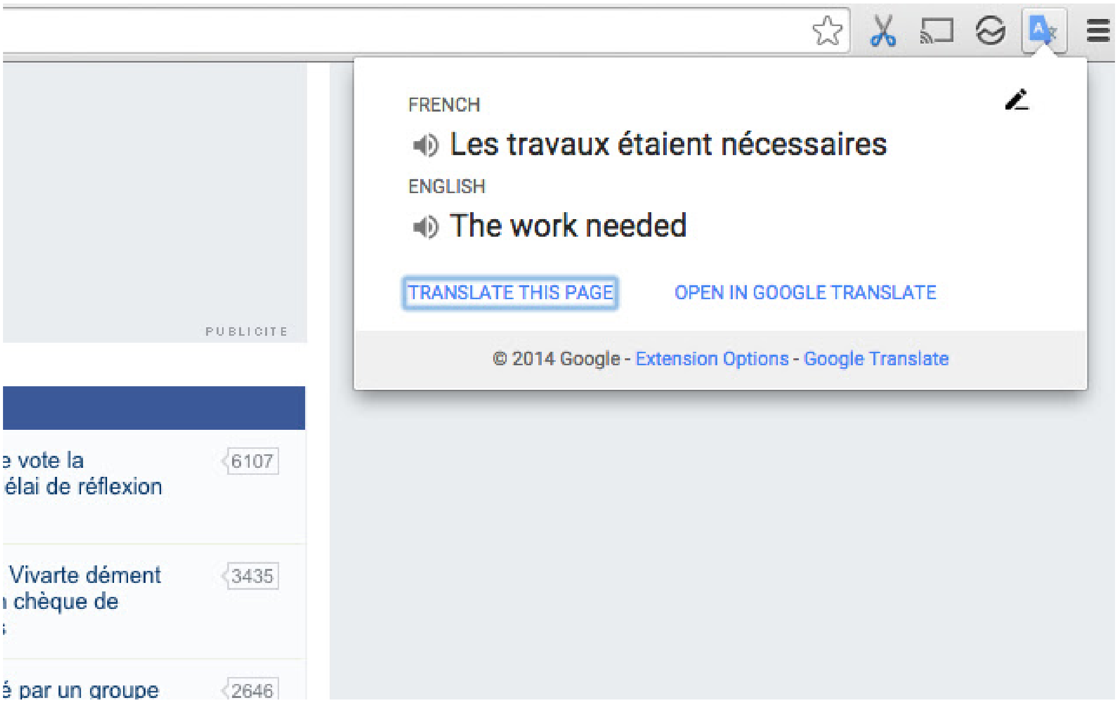

Figure 6.3: Browser Extension: Google Translate for Google Chrome provides the user with easier to access translate functionality. Fetched from `https://chrome.google.com/webstore/detail/google-translate/aapbdbdomjkkjkaonfhkkikfgjllcleb?hl=no` (Accessed 11/12/2021)

Pros As most of the APIs are shared between the most popular web browsers, an extension will be able to run on both Windows, OS X, and Linux. Thus, a web browser extension will comply with NFR 3. Modern web browsers are also regularly updated to stay secure, and NFR 2 will consequently be fulfilled. Further, the time required to implement the extension is estimated to be shorter than the time limit specified in NFR 1.

Cons As the browser extension is installed in the browser, it will not capture data traffic from programs other than the web browser. Thus, a browser extension will only capture a subset of the data sent from a computer. Second, the browsers have some specific APIs. Hence, a custom code base for each browser that the extension should support is required. In addition, people are believed to be less experienced with installing browser extensions than computer- and mobile applications, which can reduce the number of participants.

6.4 The Selected Approach

All three researched approaches have their pros and their cons. None of the approaches fulfill all the requirements. However, the time limit defined in NFR 1 was prioritized among the technical requirements, implying that the only solution to fulfill this requirement was a browser extension. The final implementation will be presented in Chapter 10.

6.5 Summary

This chapter presented the technical research that was conducted during the beginning of this thesis. In order to understand the initial requirements of the system, both Functional and Non-functional requirements were created. The three possible approaches examined were *Desktop Application*, *Mobile Application* and *Browser Extension.* Of the Non-functional requirements, *NFR 1: Implementation time less than three months* was decided to be the most important. Creating a browser extension was selected as the most suitable approach based on the requirements and the research conducted.

7 How to Make the Invisible Visible

This chapter presents the research conducted to understand how data can be visualized in order to make people aware of their digital behavior. As part of this research, an interview with Gerry McGovern, the author of the book *World Wide Waste: How digital is killing the planet and what to do about it* was performed.

7.1 Existing Solutions to Visualize Digital Behavior

7.1.1 Google

Google report that they are actively working to create products that improve the lives of people who use their products [65]. They have created wellbeing features for, among others, Android and YouTube. In addition, Google have developed "The Digital Wellbeing product experience toolkit"[1] to aid developers in developing solutions that aim at wellbeing.

Android For Android, Google have created what they have called *Digital Wellbeing*. The tool lets users get a daily overview of how they use their phone. It is designed to help users pause distracting apps and set daily limits per app. It also generates daily reports of their usage compared to other days [63]. An example of this is shown in Figure 7.1.

YouTube For YouTube, tools to let users understand their viewing habits have been created. They let the user explore how much time they spend on YouTube compared to previous time periods [63]. YouTube also reminds their user to take a break after watching videos constantly for a long time.

7.1.2 Apple

Like Google, Apple has also created solutions to make people aware of their digital habits, namely *Screen Time* and *Focus*.

[1]`https://wellbeing.google/for-developers/`

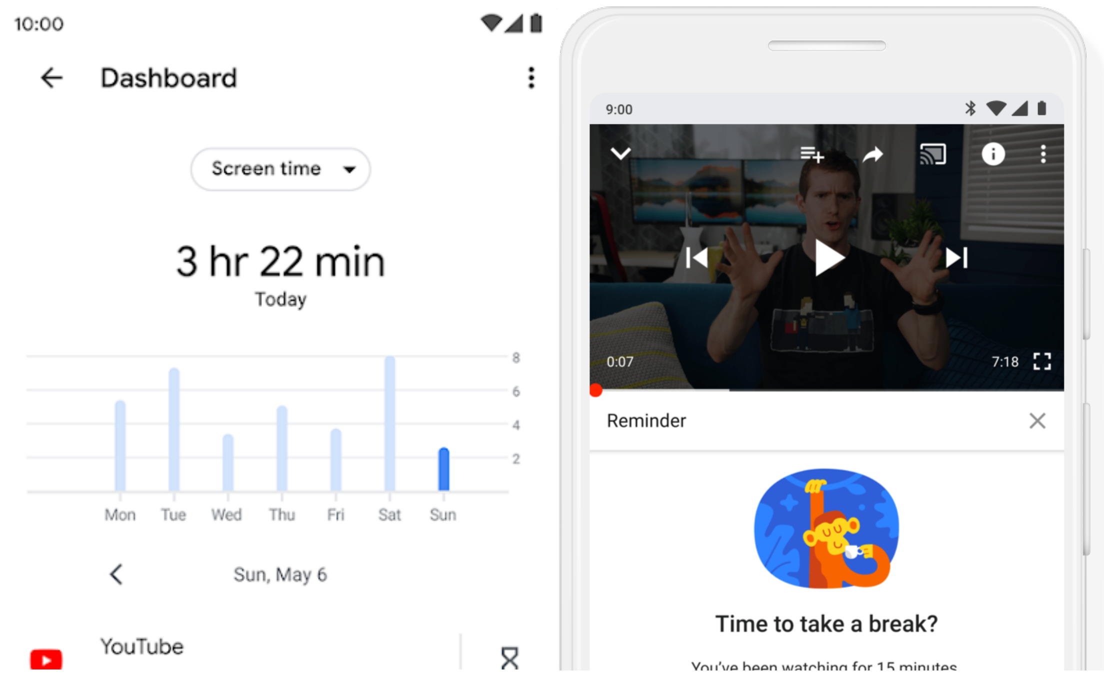

Figure 7.1: Some of Google's Wellbeing features. Fetched from `https://wellbeing.google/for-everyone/` (Accessed 11/12/2021)

Screen Time Screen Time is a feature that reports the time users have spent on their phone. It was introduced in iOS 12, to help the users better understand and control how they spend their time on apps and websites. Similarly to Digital Wellbeing, *Do not disturb mode* was introduced to help people stay in the moment. The screen time functionality generates daily and weekly reports to let users better understand their digital habits, and explore how much time they have spent on different apps. It also offers functionality to lock the apps with a code when they have been used for a given time. The Screen Time functionality is shown in Figure 7.2.

Focus Focus was introduced in iOS 15. The Focus app "lets you stay in the moment when you need to concentrate or step away from your device" [9]. It extends the functionality *Do not disturb mode* offers by allowing the user to create custom *Do not disturb mode*-modes.

7.1.3 Swedish Public Service Television Company (SVT)

SVT Play[2] is a streaming service offered by the Swedish public service television company. SVT Play had in 2020 1 370 000 active users per day, and 4 250 000 video streams were started on their platform daily. This required 1,5 petabytes of data being transferred

[2]`https://www.svtplay.se/`

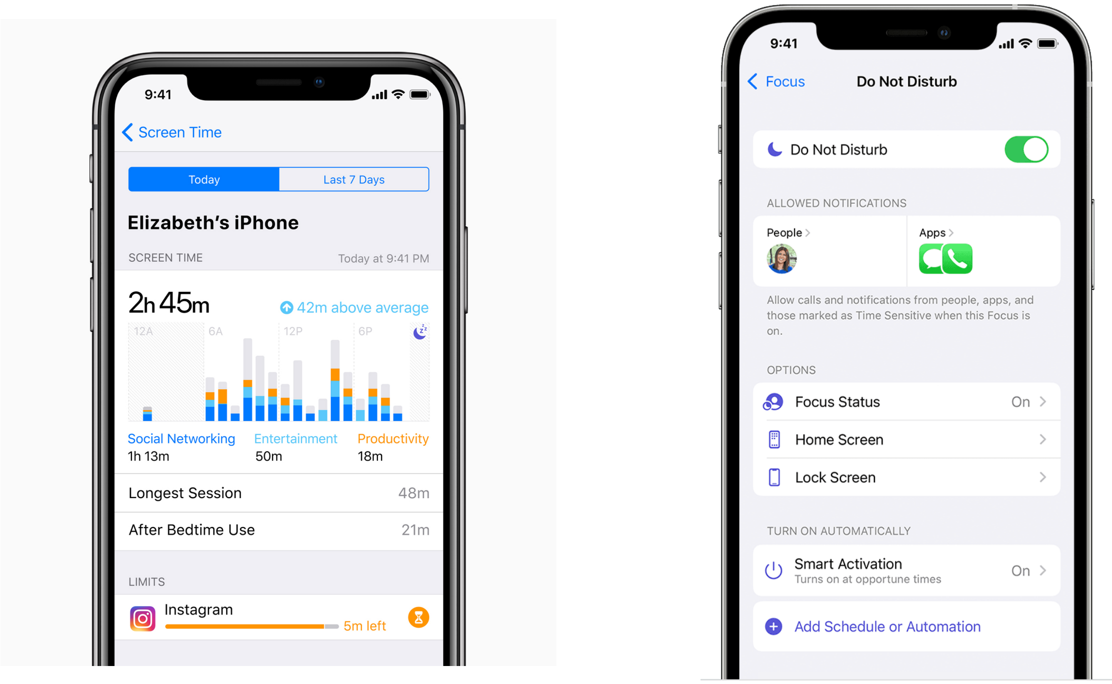

Figure 7.2: Apple's Screentime and Focus functionality. Fetched from `https://nr.apple.com/dE4l7u3E6c` and `https://support.apple.com/en-us/HT212608` (Accessed 11/12/2021)

day-to-day [131]. To reduce the amount of data being sent from their data centers, they looked at reducing the streaming quality chosen by their users. One of their suggestions was to show the user how much CO_2 the different streaming qualities would produce, as shown in Figure 7.3.

7.1.4 Oda

Oda[3] is a Norwegian online grocer. They have been working closely with the climate research institute Cicero[4] to help their customers in choosing environmentally friendly products. As a result, their customers can explore the estimated emissions of the different products Oda offers. The customers also receive a carbon receipt after shopping, estimating the carbon footprint of the purchase. After introducing the functionality, the sustainability director at Oda reported that "Our customers buy more than 50 percent more fruit and veg than the average consumer, and meat substitutes are growing 80 percent year-on-year since we added the carbon receipts" [146]. An example of a carbon receipt is shown in Figure 7.4.

[3] `https://sustainability.oda.com/`
[4] `https://cicero.oslo.no/no`

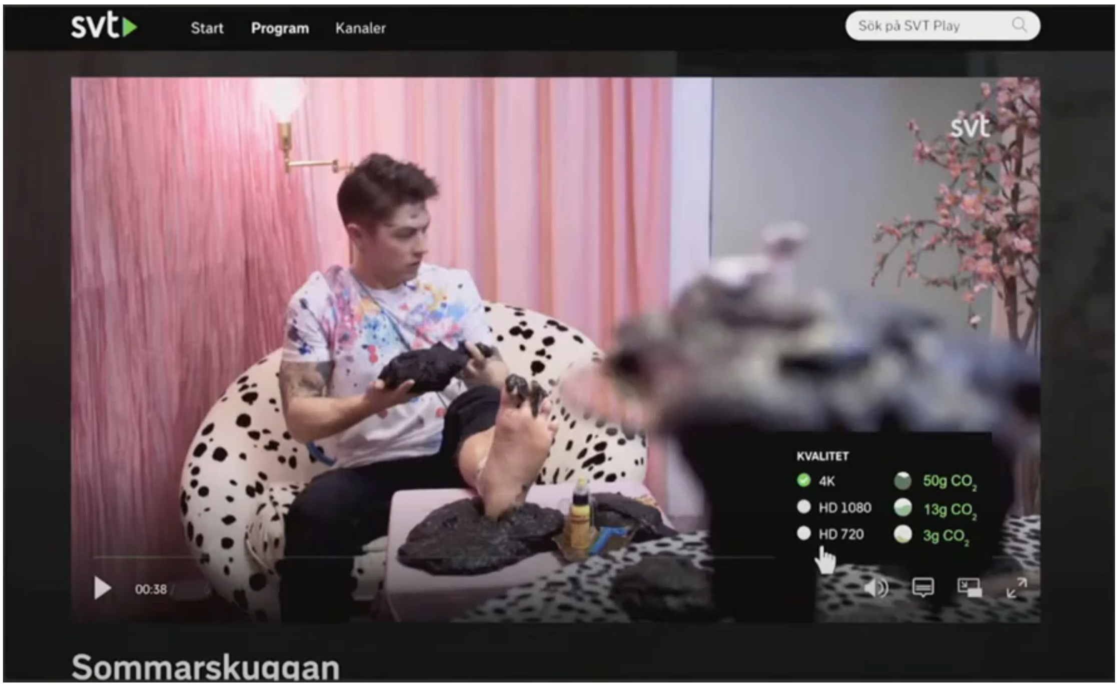

Figure 7.3: SVT Play's suggestion on how to get their users to reduce streaming quality. Screenshot from `https://youtu.be/XoOPB5i_b4Y?t=3619` (1 hour into the presentation) (Accessed 11/12/2021)

7.1.5 Hold

Hold[5] is an application created to help remove the distractions from the Internet. The application focuses on changing the students' digital habits in order to help them achieve their academic goals. The students receive points when they are not using their phones, which can be traded for physical rewards. The users can also connect to their friends and explore how much time they focus daily compared to their peers. A screenshot of the application is shown in Figure 7.5.

7.1.6 Forest

Forest[6] is, just like Hold, developed to help people stay focused. The application lets the users decide how long they want to concentrate. Motivation is triggered by growing a virtual tree. The longer they decide to stay focused, the larger the tree they will receive. If they use their phone before the time limit is out, their tree will die. Each user eventually then creates a forest of their trees, both living and dead ones. Users can connect to their friends, making it possible to focus simultaneously and explore each other forests. The Forest app can be seen in Figure 7.6.

[5] `https://www.hold.app/`
[6] `https://www.forestapp.cc/`

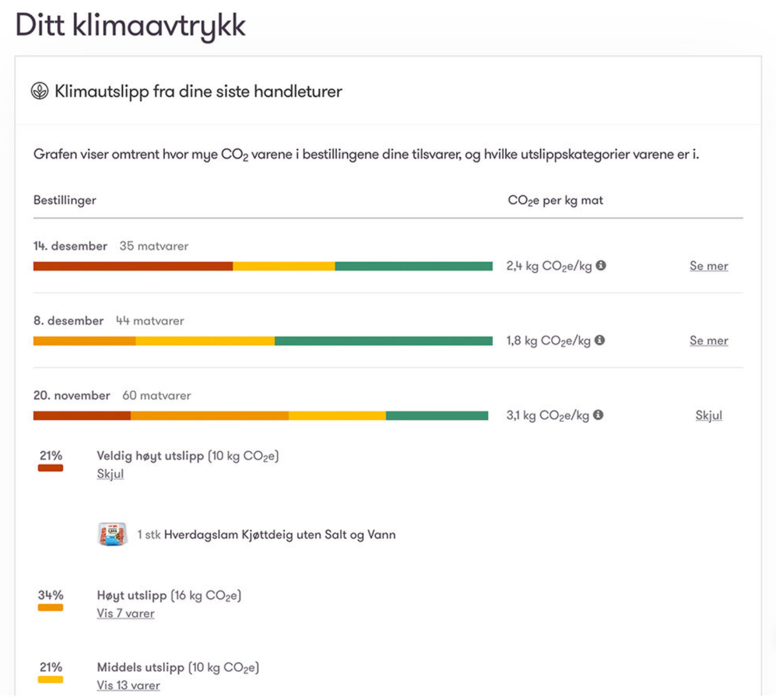

Figure 7.4: Oda's carbon receipt. Fetched from `https://sustainability.oda.com/` (Accessed 11/12/2021)

7.1.7 Website Carbon Calculator

Website Carbon Calculator[7] is a service that lets users check the carbon footprint of websites. It uses a 1.8 kWh/GB figure to estimate the carbon footprint. The amount of kWh is then converted to CO_2 equivalents based on the energy source used by the data center hosting the website. Suppose the data center is listed to run on renewable energy in the Green Web Foundation[8], they convert the energy to CO_2 equivalents using the factor of 33.4 g CO_2 equivalents per kWh. If the data center is not listed or listed as block energy, they convert the kWh using the average carbon intensity of 475 g CO_2 equivalents per kWh [39]. After entering a website to check, the user gets presented with the website's estimated emissions as shown in Figure 7.7.

[7] `https://www.websitecarbon.com/`

[8] `https://www.thegreenwebfoundation.org/`. The Green Web Foundation have tools and datasets to check if a data center runs on renewable energy.

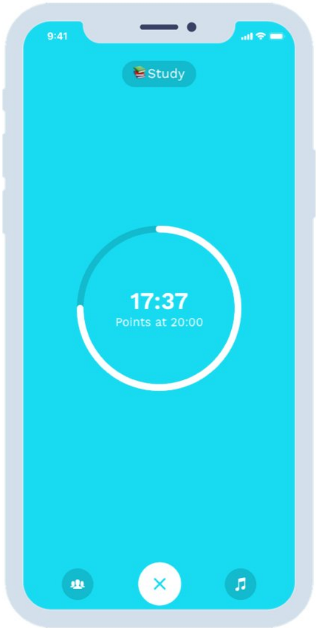

Figure 7.5: Hold. Fetched from https://www.hold.app/ (Accessed 11/12/2021)

Figure 7.6: Forest. Fetched from https://www.forestapp.cc/ (Accessed 11/12/2021)

7.2 Interview with Gerry McGovern

Gerry McGovern [9], the author of *World Wide Waste: How digital is killing the planet and what to do about it* is a highly renowned speaker [90]. He has presented how to simplify the digital experience in more than 40 countries and published eight books. He got involved in the Internet and the Web space in 1993 and started writing books about data and content management. In 2019 he became increasingly aware of the environmental impacts of the digital revolution and started to dig into the problems that digital use is causing. From an environmental point of view, he researched how the carbon footprint from digital use can be reduced. He has since then been described by the Irish Times as one of five visionaries who have had a significant impact on the development of the Web [90].

The interview with McGovern was conducted as a semi-structured interview, giving the interviewee the chance to bring up and address new issues. This was the preferred method, as the goal of the interview was to discover new themes, rather than checking

[9] https://gerrymcgovern.com/

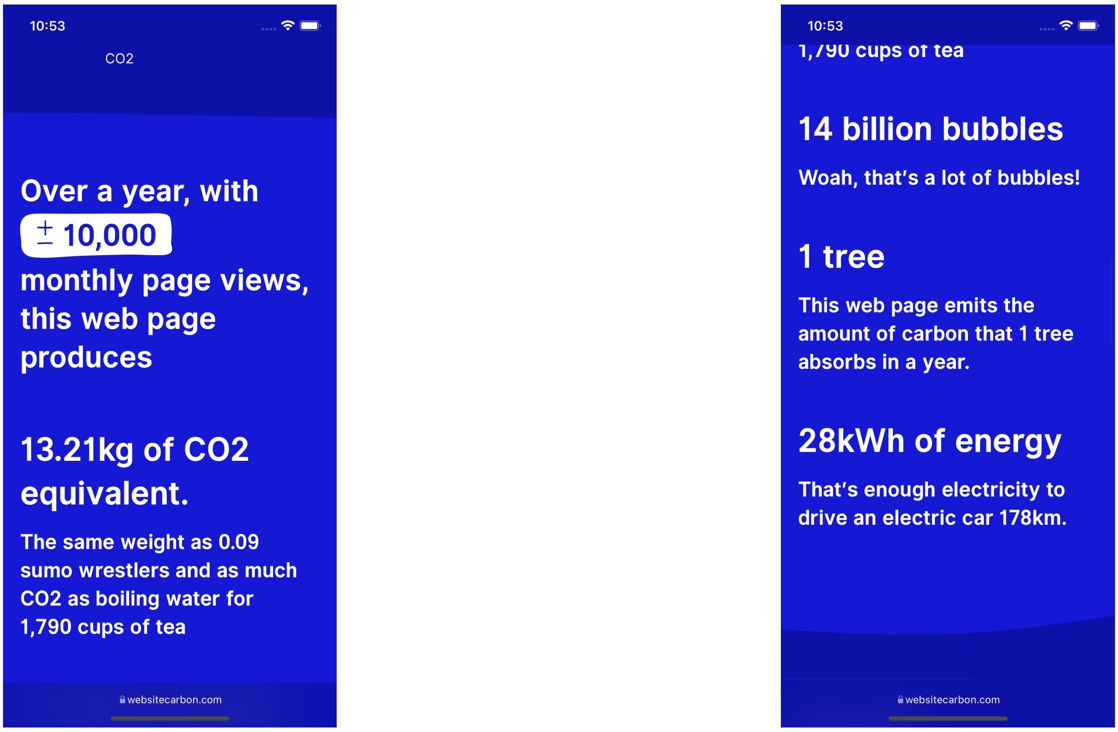

Figure 7.7: Website Carbon Calculator lets the user explore how much different websites pollute. Screenshots taken from `https://www.websitecarbon.com/` (Accessed 11/12/2021)

existing ones [109]. The interview guide is found in Appendix 1.

7.2.1 Motivation for the Interview

To my knowledge, there is no existing software that can help people become aware of their digital climate footprint. However, when I read the book by Gerry McGovern, I understood that he could provide me with valuable insights and information on how to create a system that would give users feedback on their digital carbon footprints. For the interview, I also wanted to understand his thoughts about the future of the Internet, its climate footprint, and the factors that could make it possible to reduce its carbon footprint.

7.2.2 Key Takeaways From the Interview

Information Gap The first, and most important takeaway from the Interview, is that McGovern stated that he thinks that people are unaware of the cost of the Internet. With the introduction of smartphones and the cloud, the connection between files and physical is getting vaguer. "There is a culture in which we see technology separate from energy," McGovern stated. Neither the consumers nor the producers (Developers and Designers) know how the websites they consume and produce are using energy.

Internet and its Climate Footprint. When it comes to the Internet and its climate footprint, Gerry McGovern still thinks that the study of Aslan et al. from 2015 provides the most reliable numbers when looking at the energy required for data transmission. For his book, he used the numbers of 0.015 kWh/GB based on the extrapolation of the numbers presented in the article of 2015. During his own research, McGovern found, in similarity to other researchers, that the production of the components, especially the end-user devices, is responsible for most of the pollution caused. He discovered that writing an email for 3 minutes would create 5 grams of CO_2 when using a laptop and 11 grams when using a computer with an external monitor.

Companies are Addicted to Data During his research, McGovern found that companies are addicted to data. They collect considerable amounts of data without having a clear plan on how to use it. Based on his findings, he argues that businesses typically only analyze around 10% of the data they collect. He thinks the most significant problem is that companies collect much more data than they need without organizing it, making the data potentially useless in the future.

Digital Consumer world Companies like Google and Facebook are consumer engines from a business model point of view. Their profitability is linked to consumption. Most of their income is through advertisement, which in turn drives consumption. Advertisements are often in the form of videos or images, contents of which increases the Internet's carbon footprint.

Cloud providers like Google and Apple also have the incentive to affect the amount of data their consumers create. The more data their users create, the higher the chance for the users to invest in a cloud subscription. When the amount of data a user has stored on a device reaches the storage capacity for that device, the user will most likely either (1) move the data to the cloud or (2) buy a new device. The same goes for gathering data about a user. The more data a company collects about a user, the more targeted advertisement the company can offer, resulting in a higher chance of increased profitability and consumption.

The Companies Have No Incentive to Reduce Their Digital Carbon Footprint *"As long as neither the consumers nor the governments give the companies a reason to change, they will most likely continue in the same direction"* McGovern said. He pointed out that the problem is more extensive than just data consumption; it is related to consumption in general. We need a *"Friday's for the future*[10] *kind of movement around consumption,"* according to McGovern. People need to change their consumption habits to reach the 1.5 goals. Both individuals and organizations need to consume and produce

[10] `https://fridaysforfuture.org/`

less data. He also used Apple and the Right to Repair[11] program to substantiate his argument. From his point of view, Apple would not have introduced the Right to Repair program if it had not become a social movement, making the government aware of it and forcing Apple to introduce the program.

How to Reduce the Carbon Footprint of the Internet This is a complex question. McGovern argued that the companies have the largest responsibility but will most likely not change unless the consumers either change their habits or governmental legislation forces them to change. Schools and universities are responsible for teaching people about the costs of the Internet. Developers and designers are responsible for designing and creating websites and applications that are environmentally friendly. Furthermore, we are all responsible for consuming less, both data and digital devices.

How to Visualize the Digital Carbon Footprint When visualizing the digital carbon footprint, converting the numbers into relatable units is essential. An analogy McGovern said he had been working on is visualizing behavior by converting the carbon footprint into *virtual leaves* and *trees*. The idea is then that the system visualizes the user's habits, and so the user will see a negative impact by *killing leaves* with bad habits, and when the user does something good, for instance, visiting a website on the phone instead of the computer, the user will be rewarded with *new trees*. Another vital thing to visualize is where in the world the user is polluting when using the Internet. The data centers the user communicates with are (most likely) not in their neighborhood. Most presumably, they are thousands of miles away. It is difficult for the user to understand what country they are sending data to when using the Internet. According to McGovern, showing where and how much energy and carbon equivalents different websites use around the world can change the behavior of the users.

7.3 Summary

This chapter has presented existing solutions to visualize digital habits, such as Screen Time on iOS and Digital Wellbeing on Android. Companies' solutions to reduce their consumers' carbon footprint have also been presented. Finally, an Interview with Gerry McGovern has been summarized, giving valuable insight and feedback on the process of creating an application that would make people aware of the digital carbon footprint of their Internet habits.

[11] `https://nr.apple.com/dH5n9r4q1v`

8 Requirements

This chapter will present the functional and non-functional requirements for the application. The requirements are an extension of the requirements presented in Chapter 6. They are created based on the research conducted in Chapter 5, Chapter 6, and Chapter 7. The requirements will be useful when developing the rest of the system.

8.1 Functional Requirements

A functional requirement is defined as "a requirement that pertains to a functional concern" [61]. Thus, the goal of the requirement is to define a functional aspect of the final applications. The functional requirements are shown in Table 8.1. The different requirements will be divided into smaller issues when developing the system. This will be explained further in Chapter 11.

ID	**Functional Requirement**	**Priority**
FR 1	The system should detect amount of data being sent	High
FR 2	The system should detect the location the data is being sent to	High
FR 3	The system should require authentication of the user	High
FR 4	The system should make it possible for the user to select what data that is being logged	High
FR 5	The system should update in real time	High
FR 6	The system should make it possible for the user to see what action caused what pollution	High
FR 7	The system should make it possible for the user to explore its pollution caused around the world	Medium
FR 8	The system should make it possible for users to compare pollution to the average of other users	Medium
FR 9	The system should generate reports for the user, based on its usage	Medium
FR 10	The system should provide the user with tips on how to reduce the pollution	Medium
FR 11	The system should send push notifications with usage updates	Low

Table 8.1: Functional Requirements for the system

8.2 Non-Functional Requirements

In addition to creating functional requirements, non-functional requirements are created. A non-functional requirement is defined as an "attribute of or a constraint on a system" [61]. When developing a system, it is crucial to understand the external constraints. Thus, the non-functional requirements will be of considerable aid in deciding on the system's architecture. The non-functional requirements are presented in Table 8.2.

ID	Non-Functional Requirement	Priority
NFR 1	Implementation time less than 3 months	High
NFR 2	The system should be secure	High
NFR 3	The system should be runable on Windows, Mac OS and Linux	High
NFR 4	The system should be easy to install	High
NFR 5	The system should report the data without manual intervention from the user	High
NFR 6	The system should be easy to test	Medium
NFR 7	The system should have a low carbon footprint	Medium

Table 8.2: Non-Functional Requirements for the system

9 Architecture and Potential Technology

In this chapter, the planning of the solution will be presented. The chapter will first explore the data collected by the system and how this data is used to estimate the user's digital carbon footprint. Second, the chapter will explain the actions taken to limit the carbon footprint of the stored data. Next, an overview of the planned solution will be provided, and the different components of the system will be explained. Finally, an outline of the different solutions considered for each of the architecture components will be presented.

9.1 Data to Collect

Before creating the solution, it is necessary to understand which data that needs to be collected. Collecting too much data will increase the carbon footprint of the solution and reduce the privacy of the users. Collecting too little data will potentially lead to a solution with reduced functionality, making it difficult to visualize the users' digital carbon footprint. The following data points are therefore decided as necessary:

Per user:

- Name
- Total number of network calls
- Total seconds active
- Total carbon footprint
- Total data size sent
- Total amount of kWh used

Per Network call:

- CO_2 produced
- Date (in unix)
- Domain that issued the call (Stored locally in the plugin)
- Country of the datacenter recieving the call

- Size (In bytes)
- kWh used

9.1.1 How the Carbon Footprint is Estimated

As discovered during the literature review conducted, estimating the carbon footprint from Internet usage has proven to be complicated. For this project, the carbon footprint will be estimated by first converting the size of the data sent to kWh by a factor of 0.09 kWh/GB. This approach harmonizes with the conventional approach described in Chapter 5 by converting the amount of data to energy. The estimated kWh will then be used to convert the energy to CO_2 equivalents. The factor used to convert the energy to estimated CO_2 equivalents will vary with the destination country of the network request, presented in Appendix 3.

9.2 Reducing the Carbon Footprint of the System

As defined in Research Question 3, the solution should not significantly increase the users' carbon footprint. This limits the amount of data stored per user and the number of requests sent between the different parts of the solution. Thus, the data will be accumulated to reduce the amount of data being stored. Per day, the amount of CO_2 equivalents, amount of kWh, number of bytes, and number of calls will be stored for each user. Further, to reduce the amount of data being transmitted, the data will be sent in batches, hence, reducing the number of web requests sent from the client to the server.

9.3 Overview of the Planned Architecture

Before planning the different parts of the system, the architecture is planned based on the research in Chapter 6 and 7. The architecture is implemented using a client-server architecture [22], as shown in Figure 9.1. The system will consist of the following components:

9.3.1 Browser Extension

The browser extension is the core functionality of the system. The extension is responsible for collecting the data being sent and providing the system with the attributes defined in Section 9.1. The extension consists of two separate components, communicating via event loops in a non-blocking manner.

DataCollector

The DataCollector will be responsible for collecting all the information the system needs. It will run as a background process in the browser, where it will be processing the data

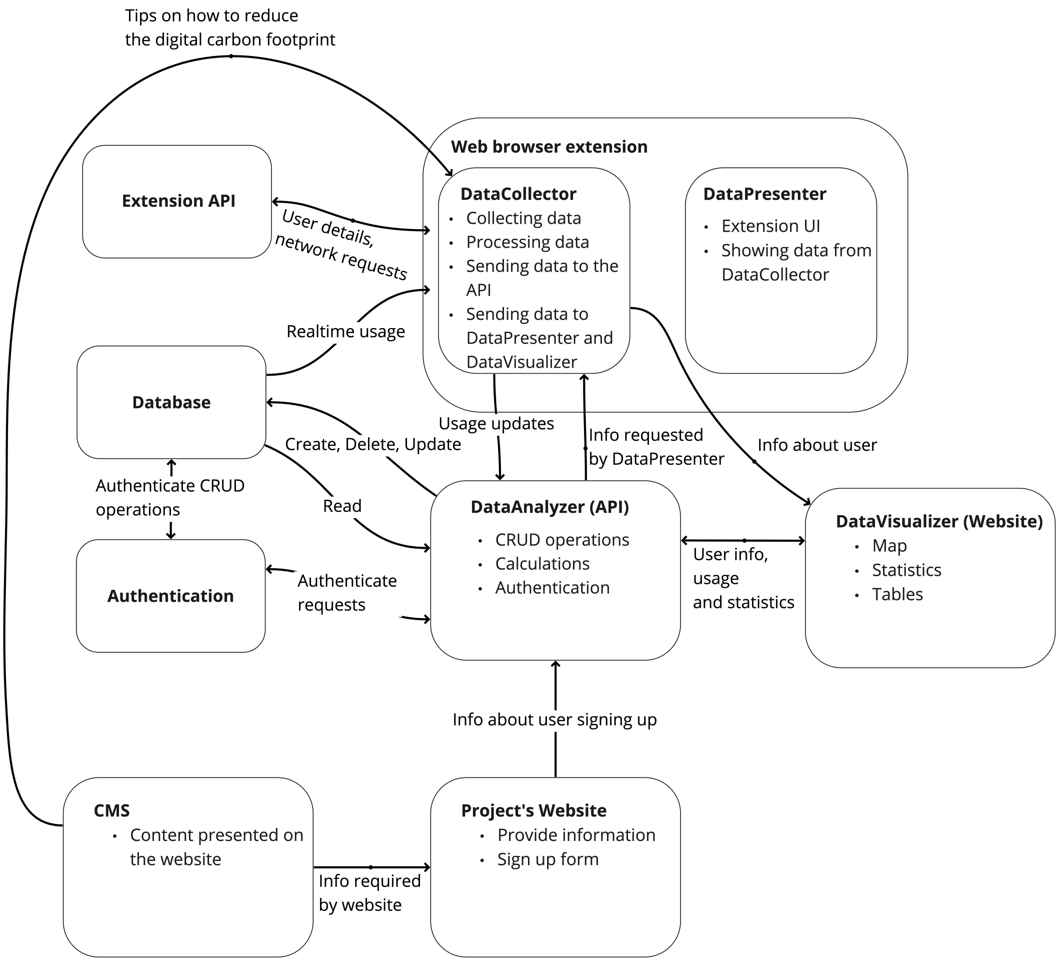

Figure 9.1: An overview of the planned architecture.

before sending the sanitized data to the API. It will also authenticate the user using the extension API of the host browser.

DataPresenter

The DataPresenter is the second component of the extension. This is the application that is shown to the user when opening the extension window. The DataPresenter will be terminated every time the user closes the application and restarted when the user opens the plugin window again. Therefore, to ensure consistency in data, the DataCollector will be responsible for handling the data. The DataPresenter will be communicating with the DataCollector through event messages to get access to the data it needs. The DataPresenter will present the estimated carbon footprint of the user in real-time and provide daily reports on the users' Internet usage. In addition, the DataPresenter will provide the user with tips on how to reduce the carbon footprint.

9.3.2 Extension API

The extension API will provide the DataCollector with the data and methods required to capture and monitor the Internet usage of a user. It will also serve the extension with the APIs required to authenticate the user.

9.3.3 DataAnalyzer

The DataAnalyzer will be the API of the system. It will be responsible for serving the different components of the application with the data they need to operate. Communication with the authentication system will ensure that only authenticated users will be able to read and store data. Based on the data reported from the DataCollector, the DataAnalyzer will estimate and update the users' carbon footprint.

9.3.4 DataVisualizer

The DataVisualizer will provide the users with more information about their usage. This will be an application where the users can explore their carbon footprint using functionality such as an interactive map and charts.

9.4 Selecting Web browser to Support

Before exploring the different solutions for the architecture's components, a decision has to be made regarding the web browser to support. As mentioned in Chapter 6, different browsers have different extensions API. Due to the limited time frame, it seems necessary to be using browser-specific APIs. Research shows that Google Chrome is the most popular browser, with between 65-85 of the market share [38, 143]. Therefore, choosing Chrome will result in more people being able to use the extension. Furthermore, Chrome is executable on OS X, Windows, and Linux, thus ensuring that Non-Functional Requirement 3 is fulfilled.

9.5 Possible Solutions for the System's Components

9.5.1 Database

When it comes to the database, three possibilities were researched. The different possible solutions for the database will now be presented:

PostgreSQL PostgreSQL[1] is "the Worlds most advanced open-source relational database" [119]. It is an open-source object-relational database, which stores the data in the

[1] `https://www.postgresql.org/`

tabular format. As the database extends the SQL language, it supports all the features of SQL. SQL benefits from standardized schemas and being ACID[2] compliant. It is optimized for large numbers of table rows, can handle large numbers of transactions in a single query, and is fast for searching and querying data [94]. On the other hand, SQL requires the schemas to be predefined. This makes it rigid, and it is demanding to change the data structure after it has been defined the first time. The architecture of SQL databases also requires them to run on only one server, so they are only vertically scalable, which can make it costly when scaling up [94, 40].

MongoDB MongoDB[3] is a popular NoSQL database used by companies like Uber and Cisco. Each record in a MongoDB database is stored as a document with key-value pair, similar to a JSON object[4]. The documents are, in turn, stored in collections. Because the data is stored in collections of documents, the data structures can be changed on the go. This makes it possible to change the data structure if the requirements of the application change. NoSQL databases are designed to scale horizontally, making scaling less costly than SQL. It is also fast for simple queries but can be slow when the queries get more complex [94]. Because there are no defined structures for the documents stored in a NoSQL database, the database cannot guarantee data consistency. Data can be malformed or missing fields, and the database architects also have to define their own rules for relations between documents. This means that documents can reference documents that do not exist [40]. MongoDB can be self-hosted or hosted through a cloud-hosted MongoDB service, called MongoDB Atlas.

Firebase Firestore Firestore[5] is a "flexible, scalable NoSQL cloud database to store and sync data for client- and server-side development" [50]. It offers much of the same functionality as MongoDB. However, Firestore was designed to connect directly to a client without an API. To be able to support this functionality, it lets the administrator define access and authentication rules for CRUD[6] operations on the different documents. It also offers real-time support, making it possible for the clients to listen to real-time updates in specific collections or for a specific document.

9.5.2 Authentication

Authentication has become an essential part of every digital solution that handles user details. Making authentication solutions is time-consuming and was outside this thesis's scope. Therefore, different authentication providers were researched to decide which one to use.

[2]ACID is an acronym that stands for Atomicity, Consistency, Isolation, Durability. These properties guarantee the correct operation of the database.

[3]`https://www.mongodb.com/`

[4]JSON Object is a way to store information. It utilizes key-value pairs and curly brackets to define objects. JSON is essentially just a string.

[5]`https://firebase.google.com/docs/firestore`

[6]CRUD is an acronym for Create, Read, Update and Delete, which are the standard database operations.

KeyCloack Keycloak[7] is "an open source identity and access management solution." [78]. KeyCloak needs to be self hosted, and supports container technology like Docker[8], Podman[9] and Kubernetes[10]. KeyCloack offer functionality for managing users and roles, supports Single Sign On (SSO) and protocols like OAuth 2.0[11]. It also supports social identity providers like Google and Twitter [153]. KeyClock is free to use under the Apache License 2.0.

auth0 Auth0[12] is an "easy to implement, adaptable authentication and authorization platform" [14]. Auth0 offers many of the same features as KeyCloack but as a service. This makes it easier to configure, as the infrastructure is hosted in the cloud by Auth0. One could also argue that it is more secure than KeyCloak, as their technology is not open-source, and they have a team of developers dedicated to ensuring the security and quality of their product. On the other side, this does not come for free; after passing 7000 active users, one has to pay based on the usage [14].

Firebase Authentication Firebase Authentication[13] is part of the Firebase platform. "Firebase Authentication aims to make building secure authentication systems easy, while improving the sign-in and onboarding experience for end-users." [51]. Firebase Authentication seamlessly integrates with Firestore Security Rules, making it easier to define access rules to Firestore. Like auth0, Firebase Authentication is offered as a cloud-hosted service. However, Firebase Authentication is free, regardless of how many users the system has [51].

9.5.3 DataCollector

The language required for a Chrome Extension is JavaScript [62]. Therefore, the only research conducted for the DataCollector is how to speed up the bundling of the application. If the application uses less time and resources when being bundled, it reduces its embodied emissions.

Parcel Parcel[14] is the "The zero configuration build tool for the web." [114]. It is used when building and bundling web projects, and offers a development server and hot reloading[15]. When building the project, Parcel uses tree shaking[16] to remove unused code. It also automatically include minifiers to reduce the output size of the code files

[7] `https://www.keycloak.org/`
[8] `https://www.docker.com/`
[9] `https://podman.io/`
[10] `https://kubernetes.io/`
[11] `https://oauth.net/2/`
[12] `https://auth0.com/`
[13] `https://firebase.google.com/products/auth`
[14] `https://parceljs.org/`
[15] Hot Reloading is a term used when the developer instantly can see code changes after saving. This improves the developer experience greatly by reducing the waiting time.
[16] Tree shaking is the process of removing unused code from the output code.

[114]. These features reduce the output file size, which will reduce the amount of data being transferred to the browser when installed, thereby reducing the system's climate footprint.

esbuild esbuild[17] is "an extremely fast JavaScript bundler" [45]. Their tests have discovered that their bundling process can be 100x faster than Parcel 2 (0.33s compared to 32.48s). In comparison to Parcel, esbuild offers a limited set of the features offered by Parcel. Where Parcel offers tools for the whole developer experience, esbuild mainly offers tools for the bundling of the project. Tree shaking and minification are natively supported, but other features need to be installed through plugins [45].

9.5.4 DataPresenter and DataVisualizer

Both DataPresenter and DataVisualizer will be built as Single Page Applications (SPA). They will be built using JavaScript and a JavaScript front-end framework. Thus, the same framework will be chosen for both applications. This will save time as code can be reused between the applications.

Vue Vue[18] is an open-source front end JavaScript framework [141]. The first version was built by ex-Google employee Evan You in 2017 and is today maintained by him. The project has in January 2022 404 contributors, and multiple sponsors [152]. Vue is a developer-friendly framework, with its clear separation of HTML, JavaScript, and CSS. The latest version of Vue also has great performance, compared to many of the modern front-end frameworks that exist [122].

React React[19] is an open-source front end JavaScript framework created by Facebook in 2013. It has in January 2022, 1533 contributors and is used by more than 8.6 million developers [48]. As React uses the JavaScript Syntax Extension, it is more difficult to understand for new developers. When it comes to the performance, it scores a bit lower than Vue [122].

Preact Preact[20] is a "Fast 3kB alternative to React with the same modern API" [120]. Preact was researched when looking for frameworks that would minimize the system's carbon footprint. Preact focuses on providing much of the same APIs as React but with better performance. It is difficult to spot the differences from a coding perspective as Preact also uses the JavaScript Syntax Extension. Compared to Vue, the performance is quite similar, Vue being a tiny bit faster [122].

[17] `https://esbuild.github.io/`
[18] `https://vuejs.org/`
[19] `https://reactjs.org/`
[20] `https://preactjs.com/`

9.5.5 DataAnalyzer

For the DataAnalyzer, Express[21] is chosen as backend JavaScript framework. Using JavaScript for the API is beneficial as the rest of the application would be transpiled to JavaScript. This would allow code sharing between the front-end and back-end code and thereby reduce the time spent at developing, complying with NFR 1 presented in Chapter 8.

9.6 Summary

In this chapter, the planned architecture has been presented. The different components' functionality and their relationship have been outlined. For each of the components, possible technologies have been explored. The planning and research conducted are used when implementing the system in Chapter 10.

[21] `https://expressjs.com/`

10 Implementing the Application

This chapter aims to present the final implementation of the application. First, an overview of the architecture is given, and the technology used to facilitate the development is introduced. Second, the database structure is explained using an Entity-Relationship Model. Third, the chapter will explore the authentication system used before each component is presented with its respective technology. Fourth, the user interface's design, the application's first iteration, and the feedback from user testing are described. Lastly, the chapter will show the final implementation of the system's user interfaces.

10.1 Code base

The code for this project is available here: `https://github.com/Jesperpaulsen/data-collector`.

10.2 Architecture Overview

The final architecture of the system is presented in Figure 10.1. Each of the components and their respective technology and relationship is shown. This figure will aid when going through how the different components were created during the project.

10.2.1 TypeScript

All the components in this project are JavaScript based. In order to reduce the chance of introducing bugs, TypeScript[1] is chosen as the development language. TypeScript is "JavaScript with syntax for types" [137]. Thus, TypeScript makes it easier for the IDE to understand the code and catch code bugs. TypeScript is converted to JavaScript during buildtime, making it runnable on every platform JavaScript runs on.

10.2.2 Linting and Formatting

In addition to the static program analysis provided by TypeScript, Prettier[2] is used for formatting and ESlint[3] is used for linting. Prettier ensures that all code files follow the same structure and increases the overall readability of the code. By the use of ESLint as

[1] `https://www.typescriptlang.org/`
[2] `https://prettier.io/`
[3] `https://eslint.org/`

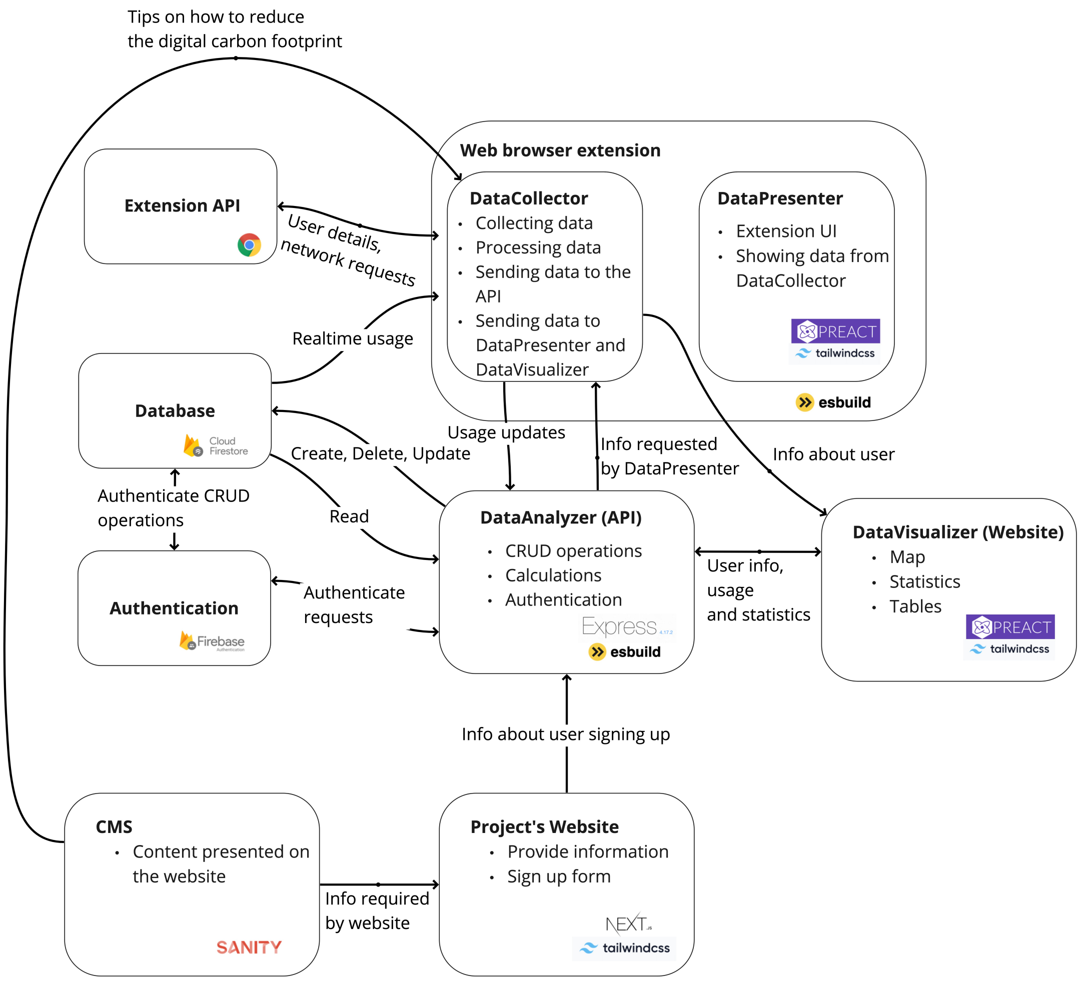

Figure 10.1: An overview of the architecture with the technology used.

linter, the IDE finds code that could introduce unwanted side effects or code that do not follow a certain style guideline [46].

10.2.3 Version Control System, Repositories and GitHub

Version Control System The Version Control System chosen for this project is Git[4]. Git is the only version control system I have experience with and is in 2022 the most used version control system [123].

Repositories The project is divided into two repositories. Both the repositories are structured as mono repositories. The reasons for using mono repositories are that some

[4]https://git-scm.com/

of the components depend heavily on each other. Therefore, using a mono repository will make sharing code between the components easier. The first repository contains the code for the DataPresenter, the DataCollector, the DataAnalyzer, and the DataVisualizer. The second repository contains the code for the Website and the CMS.

GitHub The repositories are hosted on GitHub[5]. As GitHub is the largest and most popular development platform[60], it is chosen to host the code for this project.

10.2.4 pnpm

When developing the components, external packages will be used. pnpm[6] is selected as the package manager for this project. Based on experience using the most popular package managers, such as npm[7] and Yarn[8], pnpm was chosen as a more efficient solution. "When using npm or Yarn, if you have 100 projects using a dependency, you will have 100 copies of that dependency saved on disk" [118]. Instead of storing each package inside each project, pnpm stores the plugins in a central package repository. This means that the amount of data used to download packages is reduced, which complies with *NFR 7: The system should have a low carbon footprint* presented in Chapter 8.

10.3 Database

Considering the time limit given in Non-Functional Requirement presented in Chapter 8, Firestore seemed to be the best fit for a database for this project. The Entity-Relationship model of the database is shown in Figure 10.2. The names of the different collections are defined in bold. The unique identifier (UID) of the different documents in the collection is then defined and the different attributes. Each document's document ID (UID) is stored as an attribute at the document, making it easier to access it.

As seen in Figure 10.2 there is some duplication of data. In a NoSQL database, duplication of data should not be considered a problem [132]. Even though the cost of writing and storing documents will go up with duplication of data, the read-time will be significantly reduced. On one side, this will increase the carbon footprint of the final solution, as more data will be stored. On the other side, this will reduce the number of calculations and joins required when reading from the database, reducing the solution's carbon footprint. This is beneficial when making a real-time system, as the time used to get the data to present to the end-user will not require any aggregation or joins of data.

10.3.1 Description of Some of the Attributes

In the majority of the collections, some attributes are repeated. These are used to aggregate information about the users' usage.

[5]`https://github.com/`
[6]`https://pnpm.io/`
[7]`https://github.com/npm/cli`
[8]`https://yarnpkg.com/`

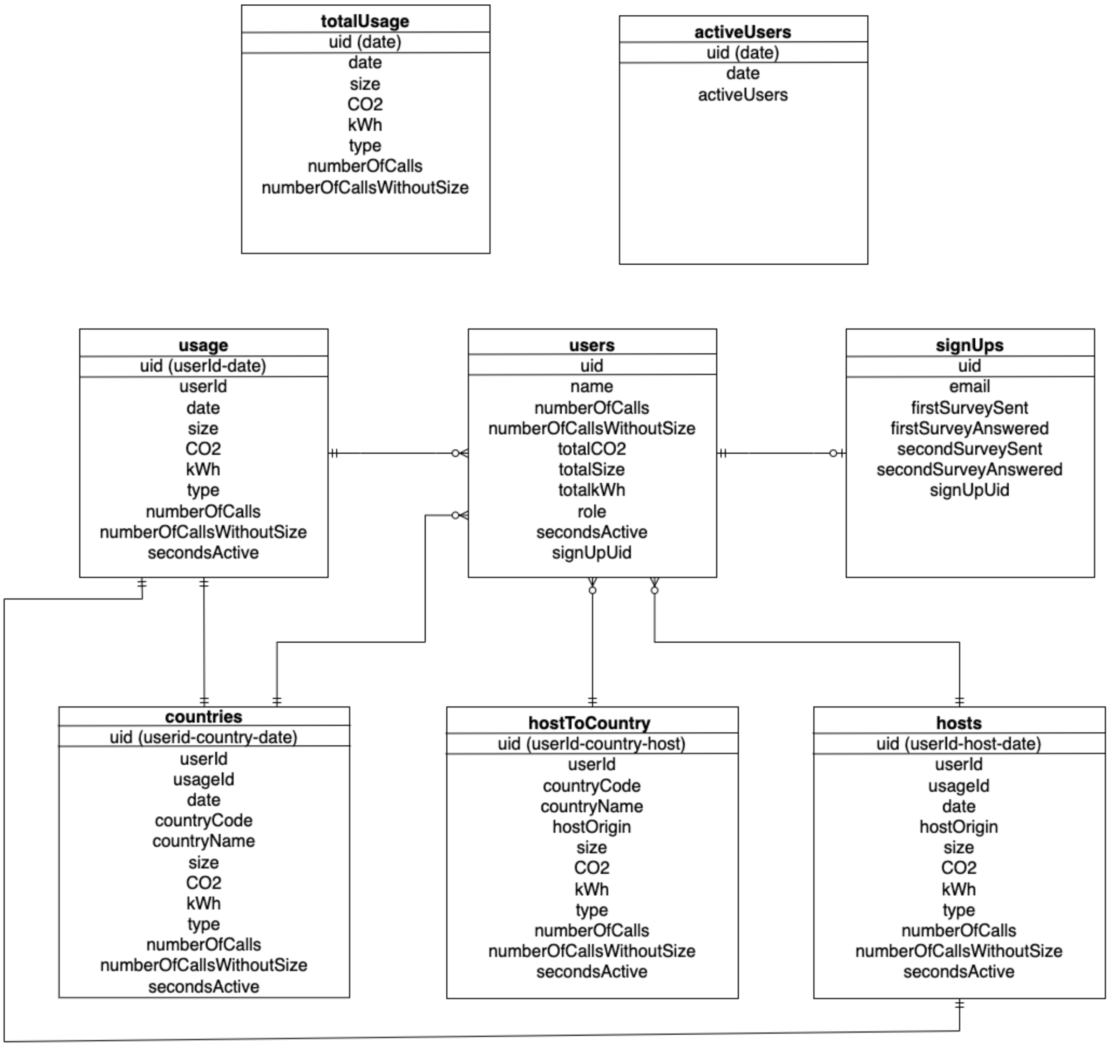

Figure 10.2: Entity-Relationship model of the database.

- **numberOfCalls**: Counter used to count the number of network calls conducted.
- **numberOfCallsWithoutSize**: Counter used to count the number of network calls where the system could not detect the size of the network call.
- **CO2 (or totalCO2)**: Counter used to estimate the amount of CO_2e emissions in grams.
- **kWh (or totalkWh)**: Counter used to estimate the amount of energy used
- **size (or totalSize)**: Counter used to measure the size of the requests in bytes
- **secondsActive**: Counter used to count the number of seconds a user has been

active.

10.3.2 Description of the Database's Collections

signUps and users The signUps collection is used when a user signs up to the project via the website. The user is signing up with their email. The rest of the attributes are used to store the time when the user answers the different surveys.

The user collection is used to store a document for each user. When a user logs in, a document in the collection is created unless it already exists. This document is given a 28 long string as UID, based on the UID generated by the authentication service. When creating this document, the signUps collection is queried to find a signUp document containing the same email. If it exists, the UID of this document is set as the signUpUid on the user document. The attributes *numberOfCalls*, *numberOfCallsWithoutSize*, *totalCO2*, *totalSize*, *totalkWh* is aggregated every time a user has been using the Internet.

usage, countries, hosts The usage, countries, and hosts collections are used to store information about a user's daily usage. Each document follows a pattern for document ID, as shown in Figure 10.2. Some of the attributes used for these collections are:

- **date**: The date attribute is stored as unix[9], and is the time 00:00 in UTC for the date the document is created.
- **countryCode**: Is the ISO 3166-1 alpha-2 code of the destination country of a network request.
- **countryName**: Is the name of the destination country of a network request.
- **hostOrigin**: Is the URL of a domain that issued a network request. It is important to mention that the hostOrigin attribute does not store the actual URL the user visited. Instead, this is stored locally in the DataCollector using a connection key. The connection key is stored in the database.

By using Firestores merge functionality, the documents are not read before the data is written. Further, by using Firestore's FieldValue[10] methods, the data is automatically aggregated by Firestore, without having to read the document. This reduces the number of network requests.

hostToCountry The hostToCountry collection contains documents that links hosts with countries. These documents are not created daily to reduce the amount of data stored. Instead, they contain aggregated data per user for the whole experiment. These documents make it possible for the user to understand what country a website is sending requests to and the estimated carbon footprint of the websites in those countries.

[9] The unix timestamp is the number of seconds since 00:00, January 1st, 1970 as UTC.

[10] `https://firebase.google.com/docs/reference/js/v8/firebase.firestore.FieldValue#static-increment`

totalUsage and activeUsers The totalUsage collection contains aggregated data for all the users for each day. The activeUsers collection contains a list of all the users that have been active per day.

10.3.3 Security

As mentioned in Chapter 9, Firestore supports writing security rules to limit access to the different collections. These are implemented to increase the privacy of the project. The security rules are presented in Appendix 2. Introducing security rules increases the number of reads of documents, but *NFR 2: The system should be secure* is seen as more important than *NFR 7: The system should have a low carbon footprint* and is therefore prioritized.

10.4 Authentication

For authentication, Firebase Authentication is selected. The reasons for choosing Firebase Authentication are that it supports all the features needed for this project, is free for unlimited users, and integrates seamlessly with Firestore. Firebase Authentication uses token-based authentication, following the OAuth 2.0 standard[11]. The tokens obey the JSON Web Token (JWT) standard[12], with information about the user embodied in the access token. Firebase Authentication also supports adding custom claims to the token for storing information such as roles. For this project two type of roles are created; *user* and *admin*. The admin has privileged access to CRUD operations on all collections in the database.

As Google Chrome is being used as the host browser for the extension, authentication is done using Google as Authentication provider, following the OpenID Connect standard[13]. This means that the users log in via their Google Account, and the user's details are provided to the API.

10.5 Firebase Local Emulator Suite

Firebase Local Emulator Suite[14] is part of the Firebase platform. It replicates many of the features provided by Firebase, such as Authentication and Firestore, and is meant to be used when building and testing applications locally [53]. By the use of the Emulator, the data used in testing and development will not affect the production data. The Emulator Suite is used in this project during the initial development and the continuous integration workflows.

[11] https://oauth.net/2/
[12] https://jwt.io/
[13] https://openid.net/
[14] https://firebase.google.com/docs/emulator-suite

Initial Development During the initial development, the emulator suite is used as an aid when scaffolding out the database structure presented in Figure 10.2. When using the emulator, creating and deleting documents and accounts is straightforward, making it possible to set up the database with mock data quickly. This is then used when writing the API using the Test Driven Development (TDD) methodology. Test-Driven Development is a coding methodology where each new feature goes through three phases. The first phase is writing the test and seeing that it fails, called the red phase. Then the simplest code that is required to get the test to pass is written, called the green phase. Finally, the last phase is to refactor the code and make sure that all tests still pass [20].

Continuous Integration (CI) Continuous Integration is a development practice used to automatically build and test code after the code has been pushed to a repository [15]. The emulator suite is used for this project when running the automatic tests. By using the emulator suite, scenarios required by the different tests can easily be set up during the CI, ensuring that all tests pass on every deployment. Another benefit of using the local emulator suite is that it could be used in the Continous Integration pipeline.

10.6 DataCollector

DataCollector is the main component of the system. Most of the other components in the architecture either depend on data from DataCollector or send data to DataCollector. As seen in Figure 10.1, DataCollector is part of the Chrome extension. DataCollector collects the size and the IP address of each network request. It is also responsible for authenticating the user and providing DataPresenter and DataVisualizer with data.

10.6.1 Technology

When writing a plugin for Google Chrome, a set of components and files needs to be in place. The most important file is the *manifest.json* file. This file is the first file that is read by Chrome during installation. It defines what permissions the plugin requires, what files Chrome should run, and what websites the extension is executed on [62]. As the final bundle requires some modification to be installed by Chrome, esbuild is used for building the project. A custom build pipeline is created to ensure that the format of the files is correct and that the files contain the necessary code to be able to run in an extension environment.

10.6.2 Authentication

The Chrome extension API exposes a toolkit for authentication using the user logged in via the browser. First, the API prompts the user to log in to the plugin. Then, the user information and a token are provided to the plugin. The token is used when communicating with the Firebase Authentication API. Using the Identity API minimizes the time required to develop a secure solution for storing tokens.

10.6.3 Monitoring the Data Sent

As presented in Chapter 6, the most crucial part of this thesis is to create a technical solution that is able to monitor the size and destination of network requests. The Chrome extension API, does not directly offer access to these data. However, some workarounds are implemented to capture the data.

Capturing the Data A simplified flow chart of how data is captured in DataCollector is shown in Figure 10.4. Because the extension API does not expose the size of the network requests and responses, two different services are created to capture the network requests. The first service is created using the webRequest API offered by the Chrome extension API. The webRequest API fires events for each network request. The events used in this thesis are marked in red in Figure 10.3.

- **onBeforeRequest**: is the first event to fire during a web request. This event is used to calculate the size of the outgoing request. The size of the outgoing request is then used when computing the total size of the request.
- **onSendHeaders**: is the second event the extension listens to. This event allows the extension to add the size of the headers to the total size of the sent request size.
- **onResponseStarted**: is the last event the plugin listens to. This event gives the plugin access to the response headers. These headers have three purposes: First, it makes it possible to calculate the size of these headers and add them to the total size of the request. Second, the plugin examines the headers for the Content-Size header. This header is optional set by servers to indicate the size of the response data. As this header is optional, it is not guaranteed to exist, and therefore a fallback method to capture the data is needed. Third, this event provides the IP address of the request. This is sent to the DataAnalyzer to determine the IP address' country.

The fallback method used in this project consists of two scripts. These are injected into each webpage after it is loaded, using Content Scripts. Content scripts are scripts that have access to the Document Object Model (DOM) and can read and modify all the details of the web pages that the browser visits [28]. The two Content Scripts are:

- **inject-xhr**: XMLHttpRequest (XHR) objects are used on many webpages to communicate with servers, without reloading the whole page [105]. By using a content script, the extension hijacks the global XHR method offered by the browser. Every time a request is sent, the content script estimates the size of the response data object and the headers.
- **inject-fetch**: The Fetch API is similar to XHR but has a more robust API that exposes more flexible methods to the DOM [99]. As websites have started to port their network requests from XHR to fetch, it is necessary to create a script to

intercept the fetch requests. In the same way, the XHR requests are monitored, the fetch requests are monitored to estimate the size of the data recieved.

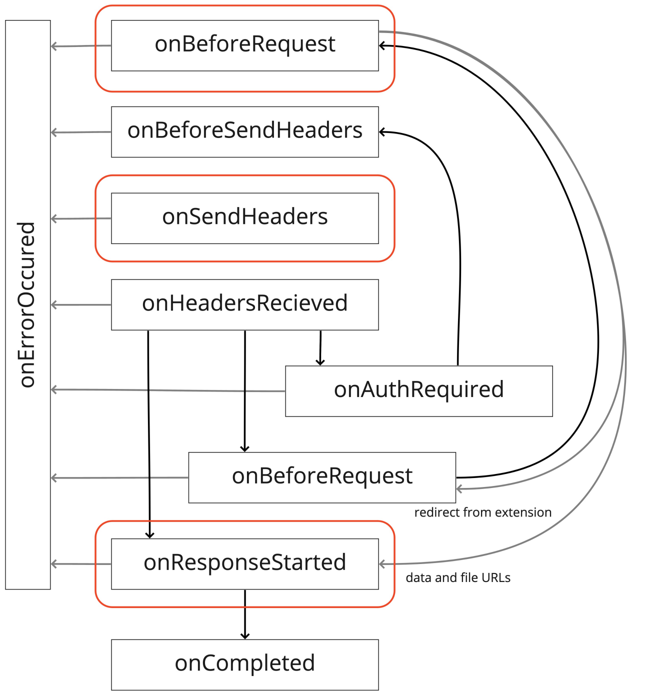

Figure 10.3: The life cycle of web request events fired by the webRequest API. The events used in this project are marked in red. Figure is reconstructed from `https://developer.chrome.com/docs/extensions/reference/webRequest/` (Accessed 11/12/2021)

.

The downside of having two different tactics to capture network calls is that there is no way to tell if a network request is counted twice. The solution to this problem is to create a *DuplicateHandler* class. This class has the responsibility of handling duplicates of network calls. Every time a network call is captured using either the content scripts or the header listeners, a timestamp is stored along with this network call. This timestamp, the URL of the webpage that issued the network call, and the URL of the target of the network call, is then hashed to a unique string, as shown in Figure 10.4. By using the hash, the DuplicateHandler can check if there exists an entry for a given request. If an entry exists, the DuplicateHandler merges the two requests into one.

Storing and Reporting Data A single web page visit can result in more than 100 unique network calls. Therefore, the system can not report every network request individually. Instead, the combination of two classes; *StorageHandler* and *DataReporter* is created to handle the data reporting. The StorageHandler has the main responsibility of storing data using the storage API offered by the Chrome Extension API. This API gives access to consistent storage, making it available between browser sessions. This means that network calls that have not been reported will be stored even if the user closes their browser.

The StorageHandler provides the DuplicateHandler with an interface to create and update network calls every time a new network request is ready to be stored or updated. The StorageHandler uses a combination of in-memory and long-lived storage. Every second, the long-lived storage is updated with the data from the in-memory storage.

For reporting the data to the DataAnalyzer, the DataReporter class is used. Every 10th second, the DataReporter fetches all the network requests from the StorageHandler. After they are transferred to the DataReporter, the requests are deleted from the StorageHandler. Next, the DataReporter sanitizes the data before sending it to the DataAnalyzer. As part of the sanitizing, the URL of the network requests issuing the network requests are anonymized. The DataReporter then sends the anonymized host, the size of the request, and the target IP address to the DataAnalyzer. If the request to the DataAnalyzer fails, the data is transferred back to the StorageHandler, and the same procedure is repeated 30 seconds later. Network calls fetched from the browser's cache or sent internally between the components of this system are not reported to the DataAnalyzer.

Privacy As privacy is a significant concern for this thesis, a class called *HostAnonymizer* is created. This class has the responsibility of anonymizing the URLs a user visited. The HostAnonymizer maintains a long-lived map of all hosts that a user has visited. Every time a new website is visited, a unique ID is generated by the HostAnonymizer. The host URL is then swapped with this ID, and the HostAnonymizer stores the updated map in the long-lived memory. If the plugin is to be deleted, this map would also be deleted, making it impossible for anyone to figure out what websites a user had visited.

10.7 DataAnalyzer

DataAnalyzer is the name of the API for this project. As seen in Figure 10.1, DataAnalyzer is responsible for most of the communication with the database, as well as the other components in the architecture. DataAnalyzer is also responsible for aggregating the user's usage attributes.

Hosting The DataAnalyzer is hosted on the Google Cloud Platform[15]. This platform is chosen because it integrates well with the Firebase Platform, and its simplicity reduces the development time. The API is deployed on App Engine[16]. App Engine is a fully managed serverless platform that handles auto-scaling, and infrastructure [32]. This means that more time can be spent on implementing the application's features.

[15] `https://cloud.google.com/`
[16] `https://cloud.google.com/appengine`

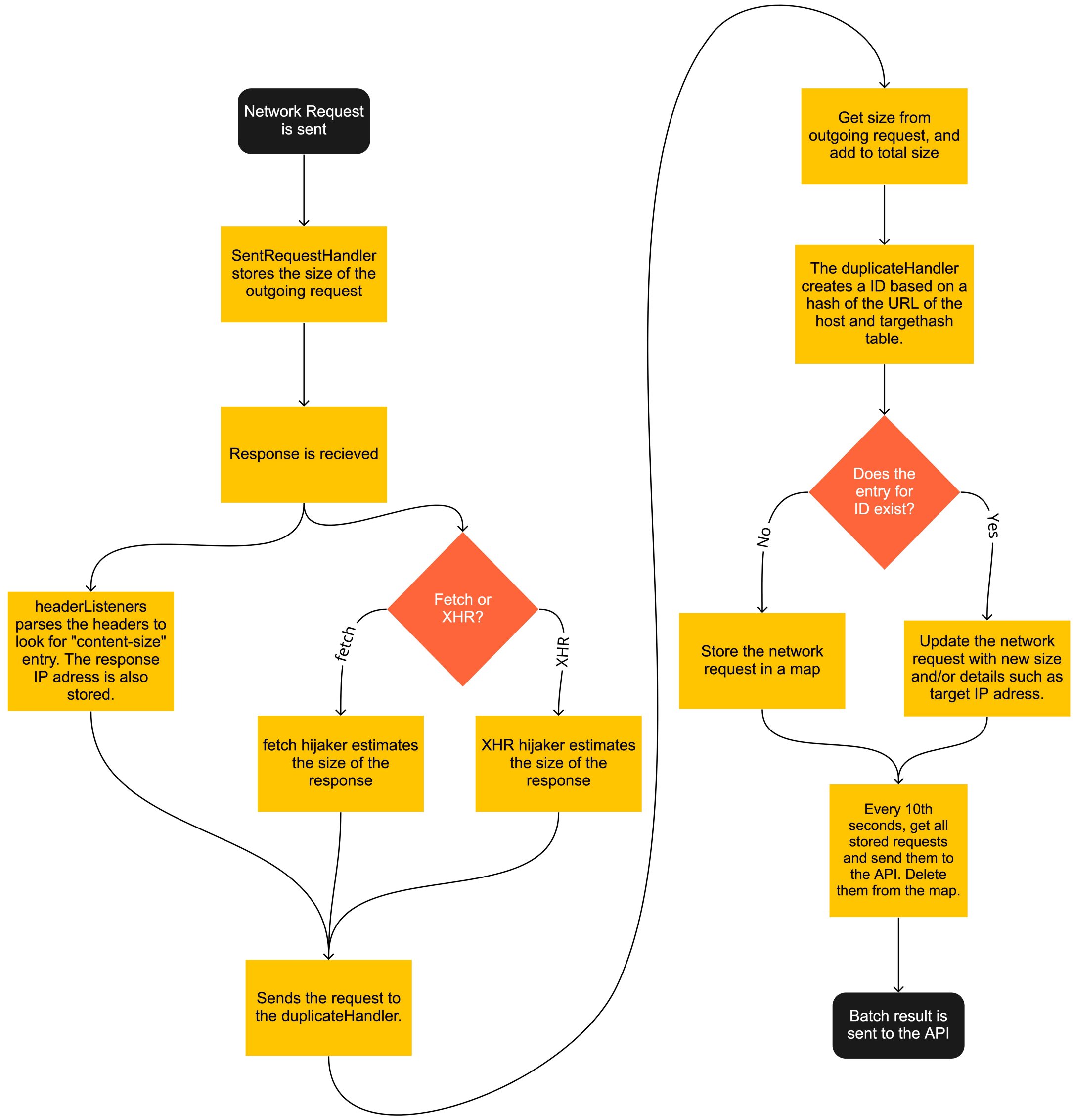

Figure 10.4: Flow chart of how data is captured and reported in DataCollector.

Carbon Intensity for Different Countries As presented in Chapter 9, the CO_2 equivalents are estimated based on the destination country of each network request. The electricity CO_2 emissions for each country are calculated based on a report called *Country specific electricity grid greenhouse gas emission factors* [55]. By creating a script that parses, and stores the results as a JSON object, where the key is the country code, and the value is specific information about the country, the emission factors for each country can be accessed during runtime. The resulting JSON object is presented in Appendix 3.

Estimating CO_2 Equivalents Per Network Request The flow for estimating the CO_2 equivalents of a network request is shown in Figure 10.5. After the network request details have been sent to the DataAnalyzer, the IP address and the size of the request are extracted. Then, the IP address is looked up using a NPM package called *geoip-country*[17]. The package maintains a local database over which IP ranges belong to what country code. By the use of this package, the destination country is extracted. Sometimes the IP address is unknown or not indexed by the *geoip-country* package; in that case, the average world carbon intensity is used. If the *geoip-country* package returns a country code, this country code is used to access the JSON dictionary created earlier to find the carbon intensity of that country. If there is no data for the carbon intensity of the given country code, the average world carbon intensity is used.

Batch Processing As mentioned, the DataCollector extension only reports data every 10th seconds. This is to reduce the number of network requests that are sent across the Internet. Instead, the network requests of a user are sent and processed in batches. The DataAnalyzer runs through every network request and updates the individual documents. It also manually aggregates the numbers that are used to update the total usage of the users to reduce the number of writes to the database.

Middleware Architecture Natively, Express is a Middleware Architecture [47]. A middleware function in Express is a method that has access to and can modify three objects; the request object, the response object, and the next middleware function. Each middleware method can also end the response. Middlewares are chained together, and the data flows through the chain. The data flow through the batch request endpoint's middlewares is shown in Figure 10.6. In addition to the middleware architecture, the DataAnalyzer is implemented using a combination of the MVC pattern and the Singleton Pattern. The MVC pattern is used on the different endpoints of the API to separate the business logic from the view [104]. The singleton pattern is used to restrict the instances of a class to only one [58]. This pattern is used when accessing the different services, for instance, the database.

Testing and Security As mentioned, the DataAnalyzer is implemented using Test Driven Development. The chosen test framework for this project is Jest[18]. Jest is a testing framework with a declarative API and little- to no-config required [74]. In addition to Jest, a NPM package called SuperTest[19] is used. SuperTest makes it easy to mock the express server, reducing the running time of the tests. The main reason for implementing tests on this project is to protect the users' privacy. Therefore, all endpoints requiring authorization are tested to check that they return the status code of 401 (Not Authorized). In addition, individual tests are written for each endpoint to make sure they handle the requests correctly. The Firestore Emulator database is also initialized with data

[17] `https://www.npmjs.com/package/geoip-country`
[18] `https://jestjs.io/`
[19] `https://www.npmjs.com/package/supertest`

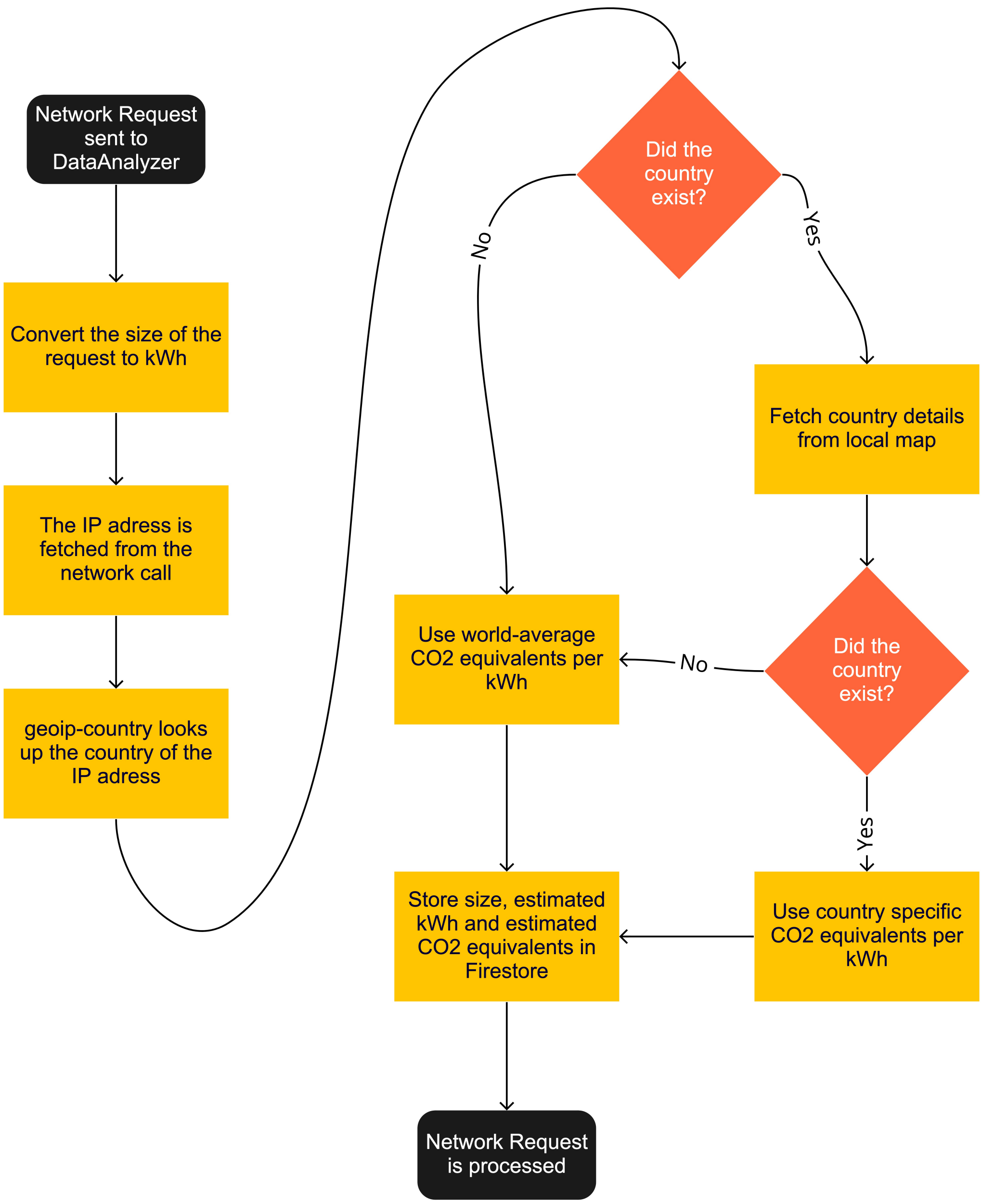

Figure 10.5: Flow chart of how the target country of a Network Request is found.

according to the different tests to ensure that no one with bad intentions can access the users' data.

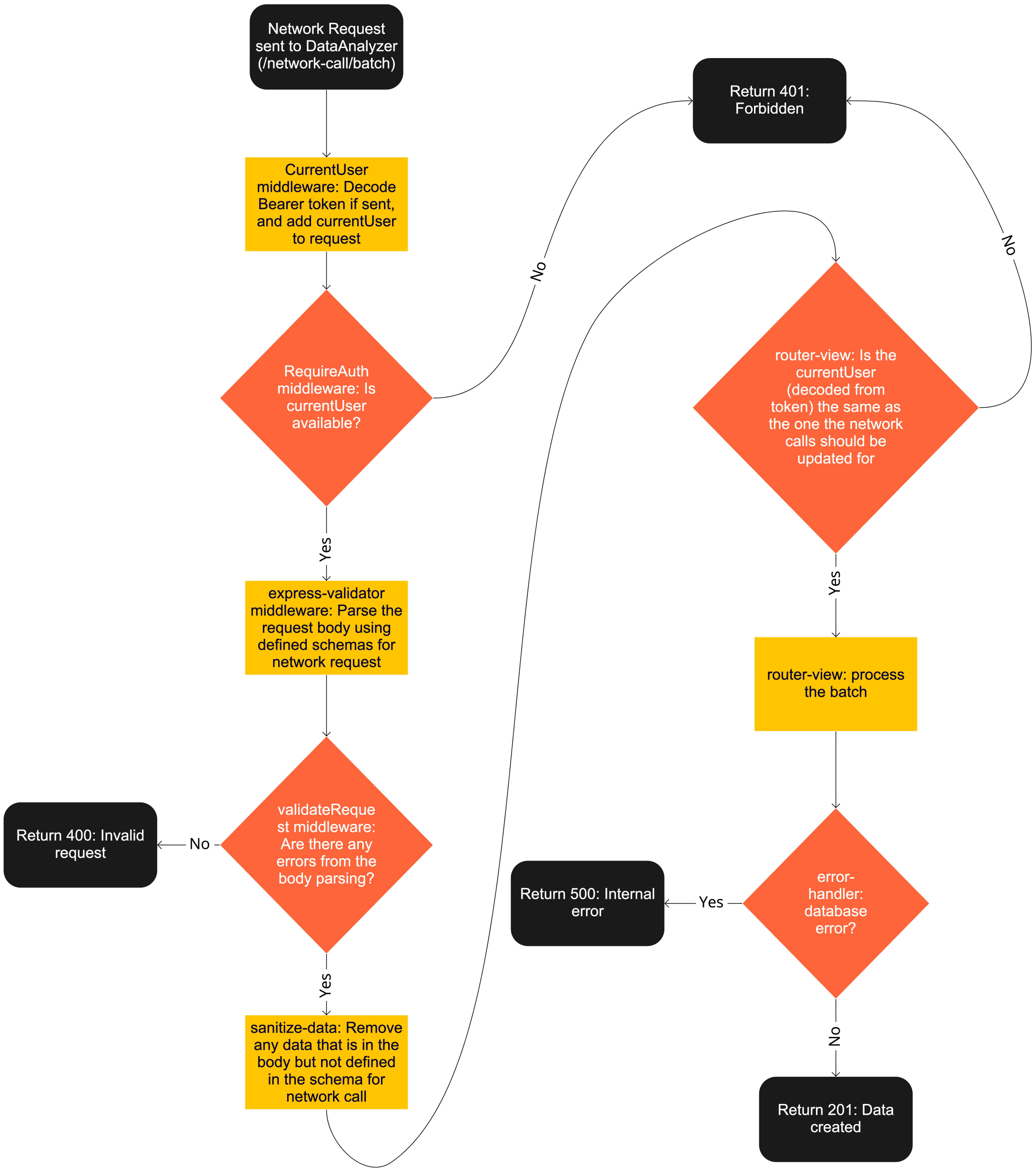

Figure 10.6: The middlewares that are used for the batch processing endpoint.

Continous Integration and Continous Delivery In order to efficiently test and deploy the DataAnalyzer after code changes, a Continuous Integration (CI) and Continuous Delivery (CD) pipeline is set up. This pipeline is run on GitHub Actions[20]. GitHub

[20] `https://docs.github.com/en/actions`

actions make it possible to automate the tests and deployment of the DataAnalyzer. On every code push to the *main* branch, the tests are executed to make sure that the core functionality of the different endpoints (including the security) is still working as intended. After the tests have passed, GitHub Actions are used to deploy the new version of the API.

10.8 Design and First Iteration of DataPresenter and DataVisualizer

Both DataPresenter and DataVisualizer are implemented using Preact. In addition to this Tailwind[21] is used for styling. Tailwind is a utility-first framework that scales regardless of the size of the project. Tailwind utilizes tree-shaking and purging to make sure that only the CSS styles used in the applications are part of the final bundle. This reduces the final build size [133].

10.8.1 Designing the Different Parts of the Applications

Figma[22] is used when designing the applications. Some of the sketches produced during the design phase are shown in Figure 10.7, Figure 10.8, Figure 10.9 and Figure 10.10.

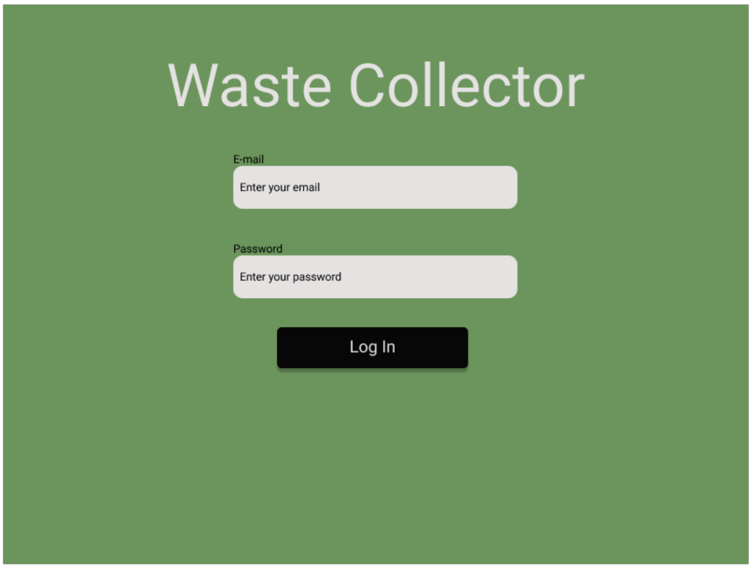

Figure 10.7: A sketch of the login screen for the extension.

[21] https://tailwindcss.com/
[22] https://www.figma.com/

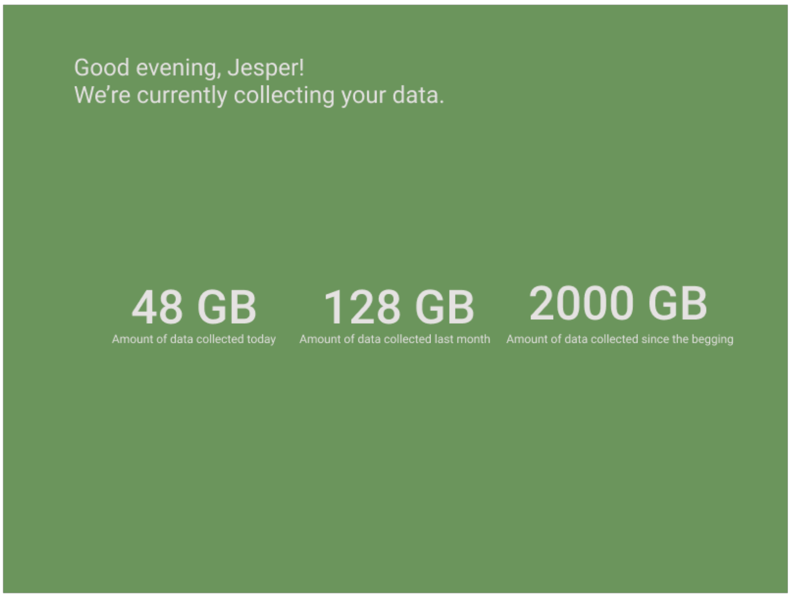

Figure 10.8: A sketch of the DataPresenter's dashboard.

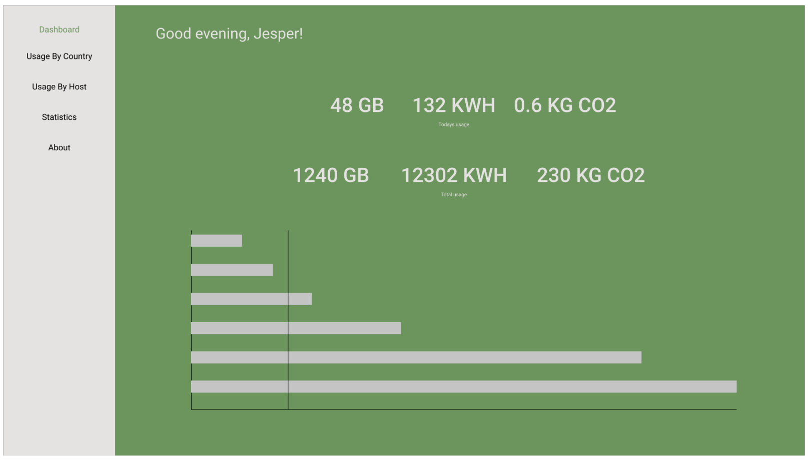

Figure 10.9: A sketch of the DataVisualizer's dashboard.

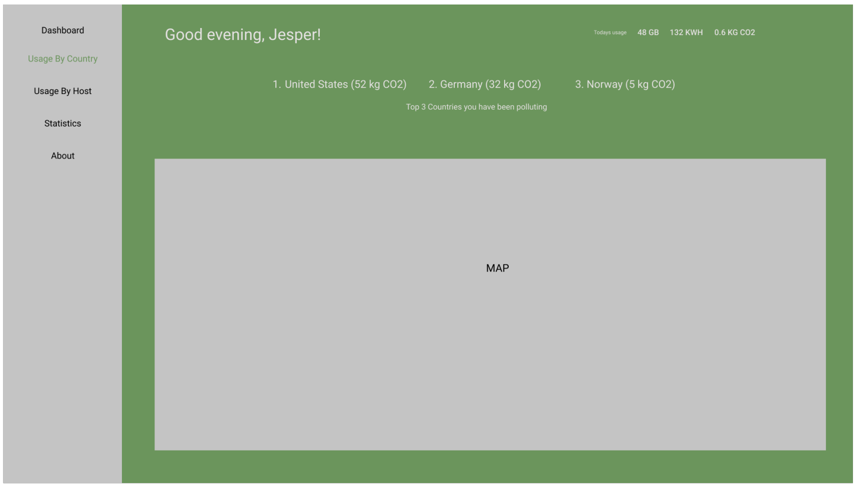

Figure 10.10: A sketch of the DataVisualizer's map functionality.

10.8.2 First Iteration

The DataPresenter and the DataVisualizer are implemented based on the sketches created. The functionality of the first iteration was limited to the Dashboard presented in Figure 10.11, a simple chart over the user's usage from the last week, and a world map shown in Figure 10.11. The application was used to conduct two user tests.

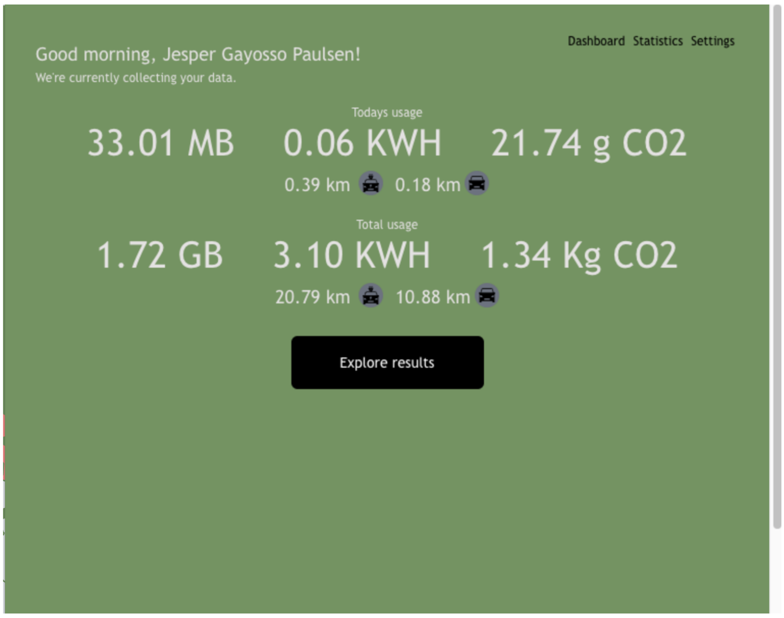

Figure 10.11: First iteration of the DataPresenter's dashboard.

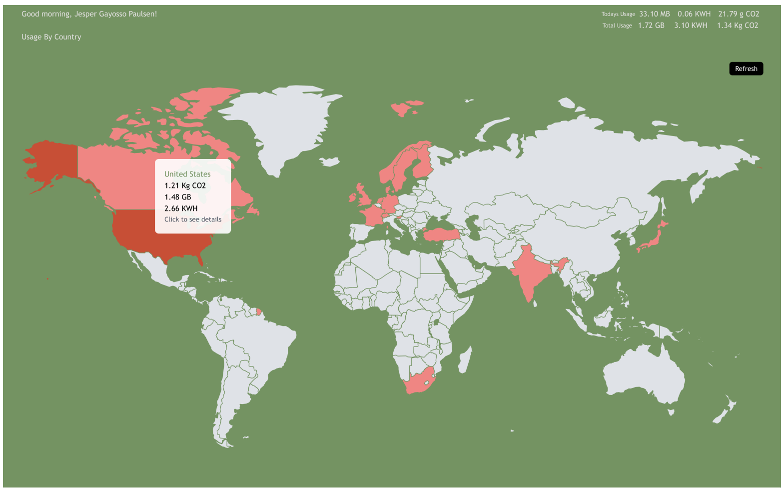

Figure 10.12: First iteration of the DataVisualizer's map functionality.

10.8.3 Feedback From User Testing

Due to the limited time available and the limitations of Covid-19, only two user tests are conducted. The user testing guide is presented in Appendix 4. During these user tests, important feedback is received. The feedback is summarized in these bullet points:

- The DataPresenter should include tips.
- The DataPresenter should include reports.
- The DataPresenter should use push notifications.
- The DataPresenter's black-listing functionality of domains should be easier to understand.
- The DataVisualizer should show what websites a user visited when causing pollution in a country
- The DataVisualizer should make it possible to select the type of usage to explore on the world map
- The DataVisualizer should make it possible to select what type of usage to explore statistics for
- The DataVisualizer should make it possible to see details about the usage each day
- The overall design of the applications should be improved (The green color should be removed).

This feedback is used to develop the final implementation.

10.9 Final Implementation of DataPresenter

After the first iteration, the final implementation of the DataPresenter is created. The design is updated, and new features are added based on the feedback from the user testing.

10.9.1 Dashboard

The final dashboard of the DataPresenter is shown in Figure 10.13. After the user testing, it became prominent that the overall design needed to be updated. Therefore, the green color from the sketches was swapped with a lighter green color. The icons for electric and petrol car is also updated. White boxes are used to make it easier for the user to understand what information is linked together.

Today's usage The design of today's usage in Figure 10.7 is influenced by the applications that were researched during Chapter 7. Forest and Hold, displaying the remaining time in real time, was a great inspiration when creating the live numbers. The way SVT Play presented the estimated emissions from streaming in different resolutions introduced the idea of showing the estimated CO_2 emissions from the user's Internet usage.

Statistics Under the display of today's usage, statistics on the user's pollution are shown. The numbers are a percentage comparison of the users' usage today compared to their own usage the last week and all the others' average usage from the last week. By comparing the users' usage to everyone else, gamification was introduced. This was implemented based on the research of Hold, to aid the participants with motivation to reduce their digital carbon footprint.

Tips and Challenges The last box contains tips and challenges on how to reduce the amount of data that is transferred. During the user testing, it became clear that the participants needed a way to understand how they could reduce their digital carbon footprint. Therefore, tips were introduced. The tips help the user understand how they can reduce their digital carbon footprint, and a new tip is fetched daily from the CMS. The tips are shown in Appendix 6.

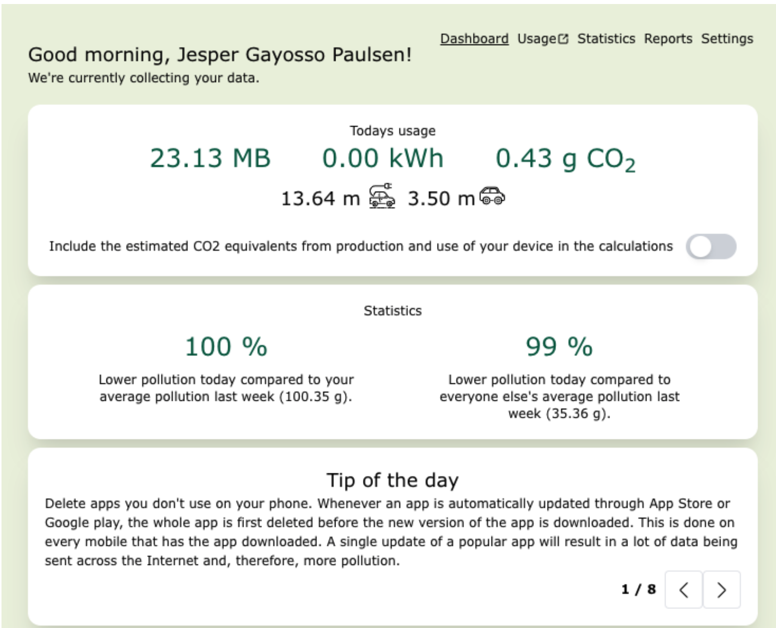

Figure 10.13: Final implementation of the DataPresenter's dashboard

10.9.2 Usage

The usage button in the menu opened the final implementation of DataVisualizer, which will be presented in Section 10.10.

10.9.3 Statistics

The statistics page is shown in Figure 10.14. This page contains two boxes similar to the ones seen in the Dashboard. While the Today's usage box is equal to the one on the Dashboard, the Total usage shows the user's total usage from they started using the extension. The statistics page also shows a graph of the user's estimated CO_2 emissions and energy usage for the last week. This is inspired by the features of Apple's screen time, Android's wellbeing, and Oda's carbon receipt, presented in Chapter 7.

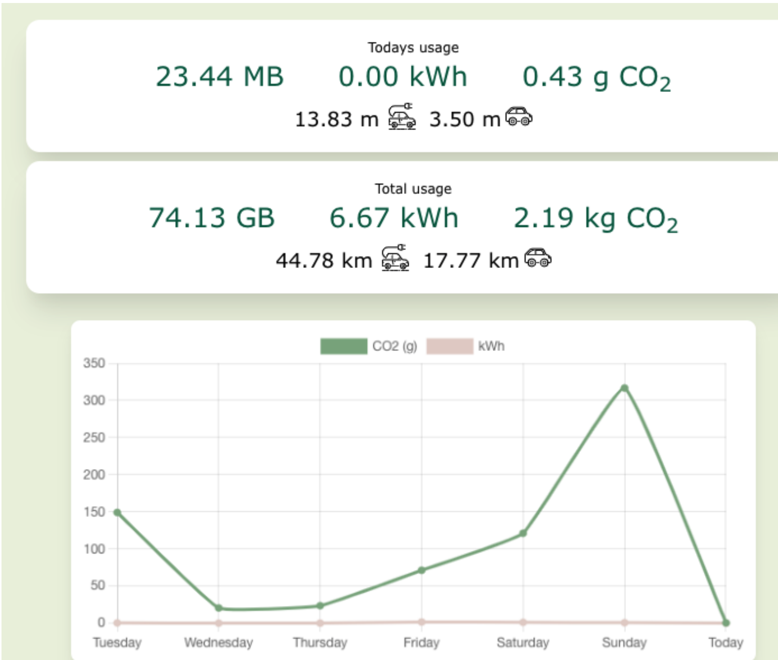

Figure 10.14: Final implementation of the DataPresenter's statistics page

10.9.4 Reports

The reports page is shown in Figure 10.15. A new report is produced every day based on the user's usage yesterday. It contains information about the user's usage that day, compared to the user's and everyone else's usage for the last week. This feature is inspired by Apple's screen time reports and Oda's carbon receipts, as well as the gamification mentioned in Hold and Forest, presented in Chapter 7.

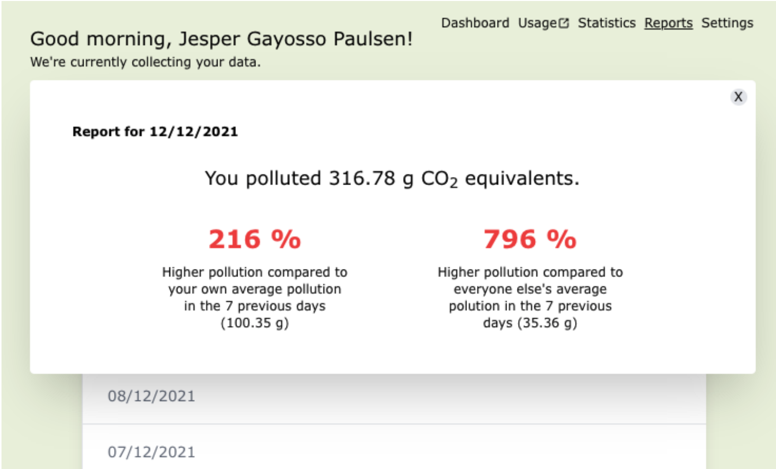

Figure 10.15: Final implementation of the DataPresenter's reports page

10.9.5 Settings

The settings page is shown in Figure 10.16. The settings page makes it possible for the user to blacklist domains from which the extension should not report usage. It also makes it possible for the user to reset the live counter if there seems to be an error.

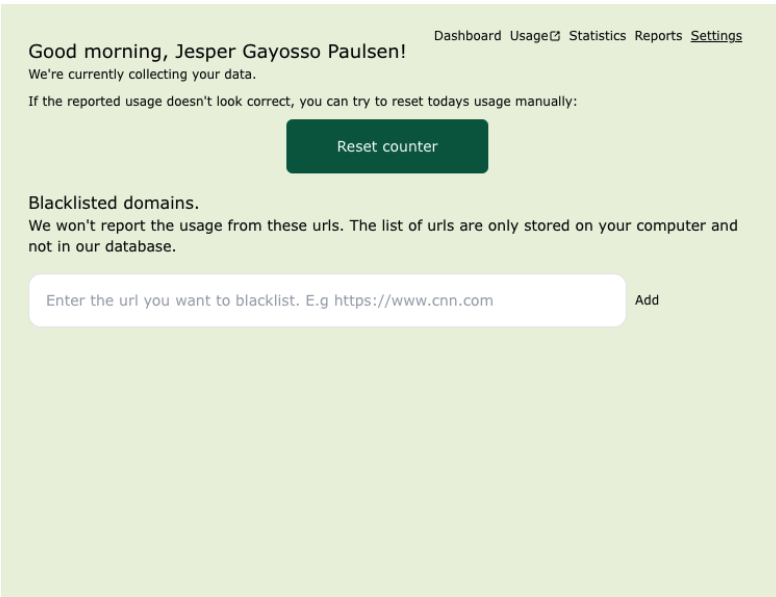

Figure 10.16: Final implementation of the DataPresenter's settings page

10.9.6 Push Notifications and Badge

For the user to be reminded about their usage, two features are introduced. The first feature is to update the badge shown in Figure 10.17. The badge shows the estimated CO_2 emissions of the user in real-time. The badge will turn red if the user uses more than their average for the last seven days. In addition to this, push notifications are sent from the DataCollector, to the web browser. An example of a push notification is shown in Figure 10.18. Every new day, a push notification containing information about the user's usage the previous day is sent. In addition to this, a warning about the user's usage is sent when the user's estimated CO_2 emissions pass the following limits per day:

- **Passing 10g and 100g CO_2 equivalents**
- **Passing their average for the last seven days**
- **Passing the average usage off all the other users for the last seven days**

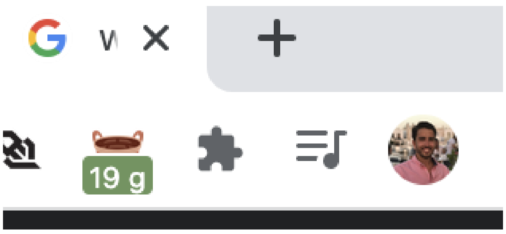

Figure 10.17: Final implementation of the DataPresenter's badge

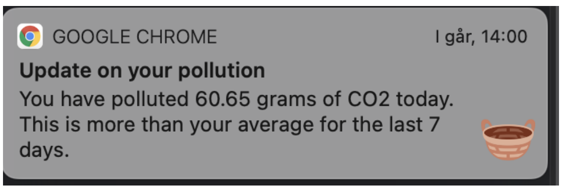

Figure 10.18: Example of the final implementation of the DataPresenter's push notifications

10.10 Final Implementation of DataVisualizer

After creating the final implementation of the DataPresenter, the DataVisualizer is finished. Some of the components from the DataPresenter are reused between the applications, such as the buttons and input fields.

10.10.1 Technical Details

In order to be able to display the actual hosts and not the anonymized ones, the DataVisualizer is dependent on talking with the DataCollector. This is done using a website to extension API offered by the extension API. When the user visits the DataVisualizer, the hosts are sent from the DataCollector to the DataVisualizer. Every new host added to the DataCollector is also sent to the DataVisualizer. Furthermore, the API sends the user details from the DataCollector to the DataVisualizer.

Hosting The DataVisualizer is hosted on Firebase Hosting[23]. Firebase Hosting is a hosting solution for static websites. Firebase Hosting was chosen as it is free to use, SSL certificates are generated for custom domains, and it makes it easy to deploy new versions of the website [52]. The DataVisualizer is accessed at `https://dashboard.jesper.no`.

10.10.2 Overview

When entering the DataVisualizer, the user is shown the overview map, presented in Figure 10.19.

Color and filter The countries a user has sent data to are marked with red. The shade of red indicates either the amount of data sent to that country, the number of requests sent to the country, estimated CO_2e emissions in the country, or the estimated energy usage in kWh in that country, relatively compared to the other countries. The attribute to show on the map is chosen using the dropdown menu on top of the map.

Hovering When hovering a country, details about the user's usage in that country is presented, as seen in Figure 10.20.

Searching The overview map also makes it possible to search for the domains a user has visited, as shown in Figure 10.20. Only the countries the selected domain has sent data to are highlighted.

Details When clicking on a country, more details such as the source used to estimate the CO_2 equivalents are shown, as presented in Figure 10.21.

[23] `https://firebase.google.com/products/hosting`

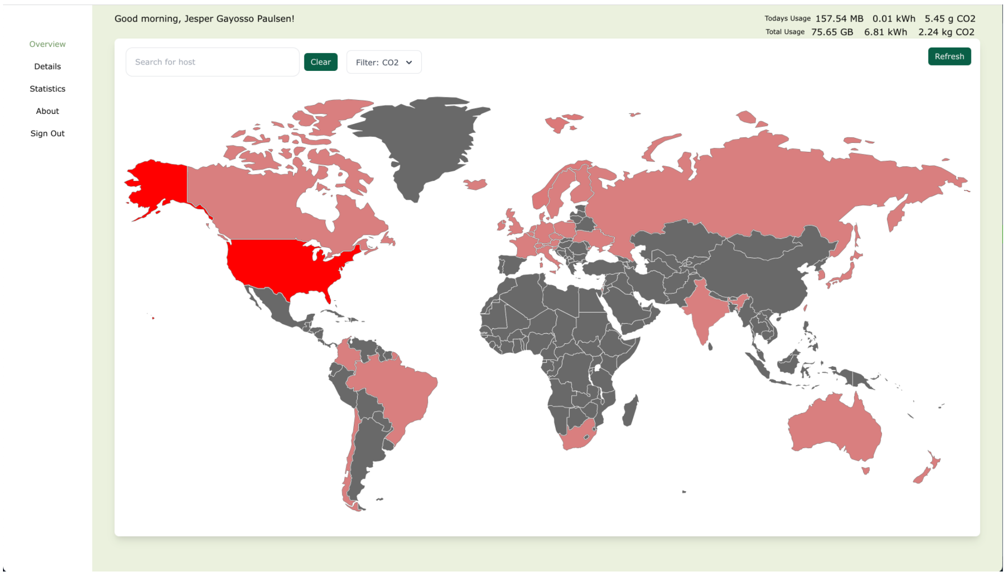

Figure 10.19: Final implementation of the DataVisualizer's map

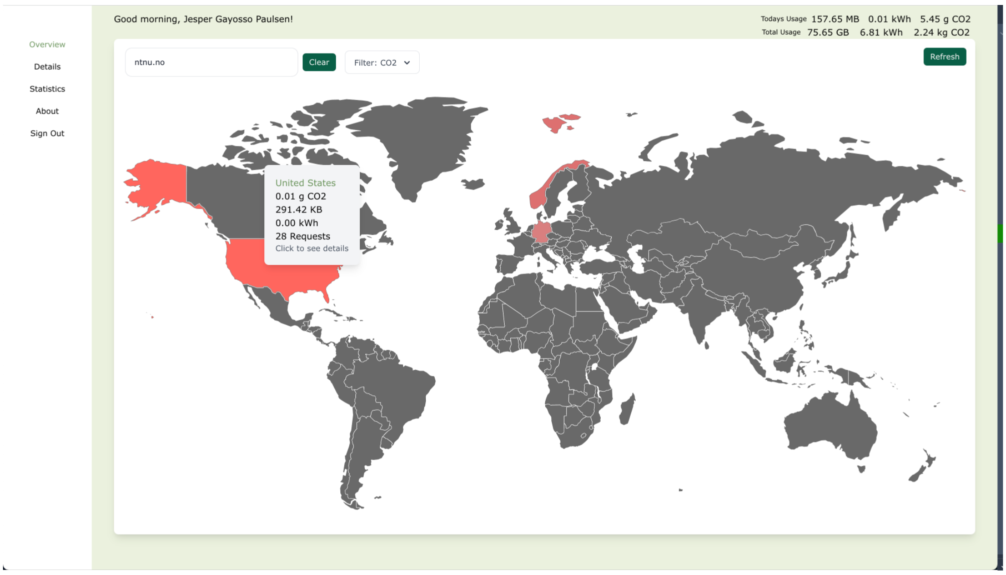

Figure 10.20: Final implementation of the DataVisualizer's map. Here the user first has selected to see info for the domain *ntnu.no*. Then the user hover the mouse over the USA and receives some info about the usage.

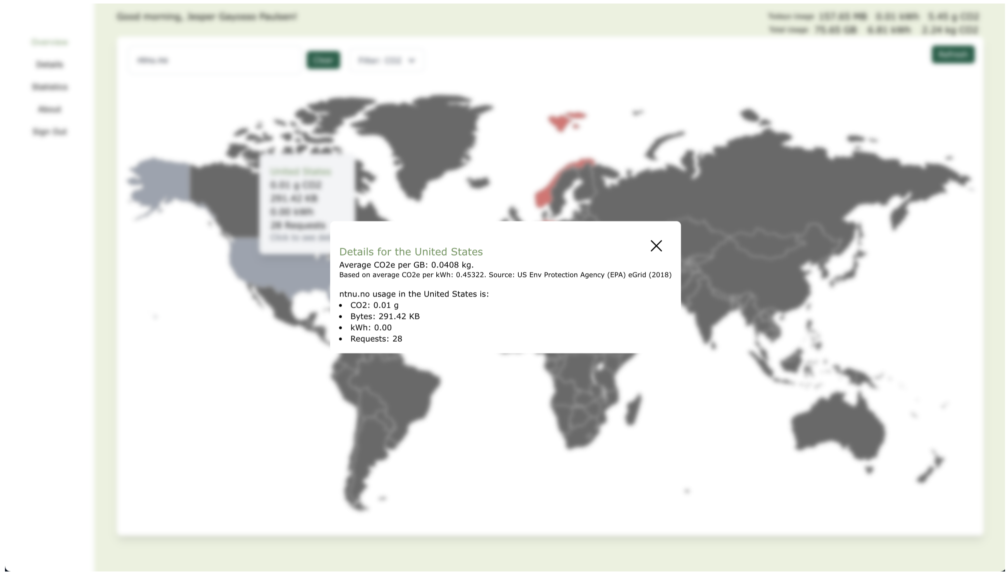

Figure 10.21: Final implementation of the DataVisualizer's map. Here the user have clicked on the USA.

10.10.3 Details

In order for the user to explore their data usage in more detail, a table containing the aggregated data for all the domains the user has visited is created, as presented in Figure 10.22.

Selecting date The date picker on top of the table makes it possible to see the data used on specific dates. This is illustrated in Figure 10.23.

Sorting Initially, the table is sorted alphabetically. However, it is possible to sort the table on any of the other attributes. In Figure 10.23, The CO2 column is sorted in descending order (Indicated with the grey background on the header column).

Searching In Figure 10.24, the search functionality of the details table is presented. The search bar makes it possible for the user to see the usage of specific domains.

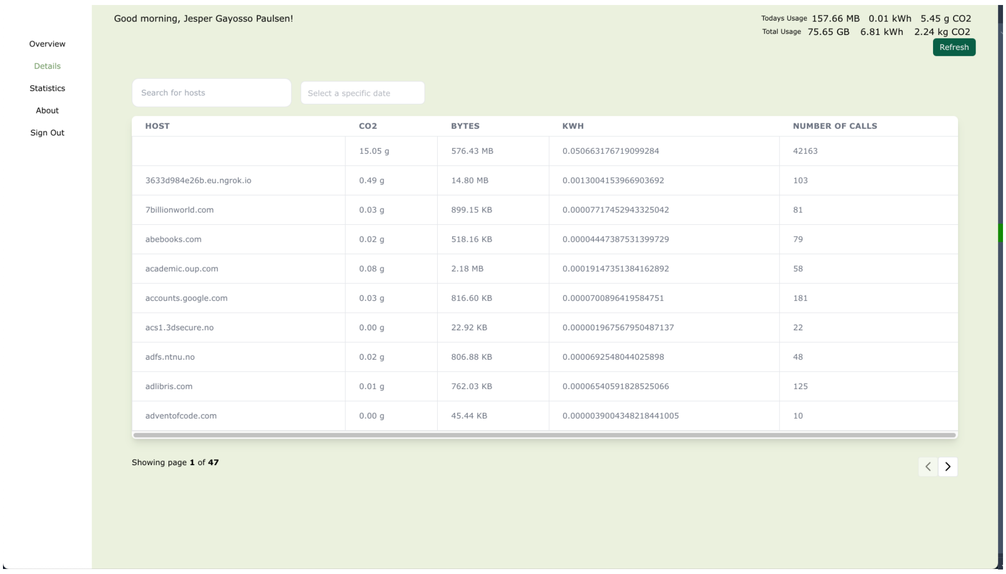

Figure 10.22: Final implementation of the DataVisualizer's details table

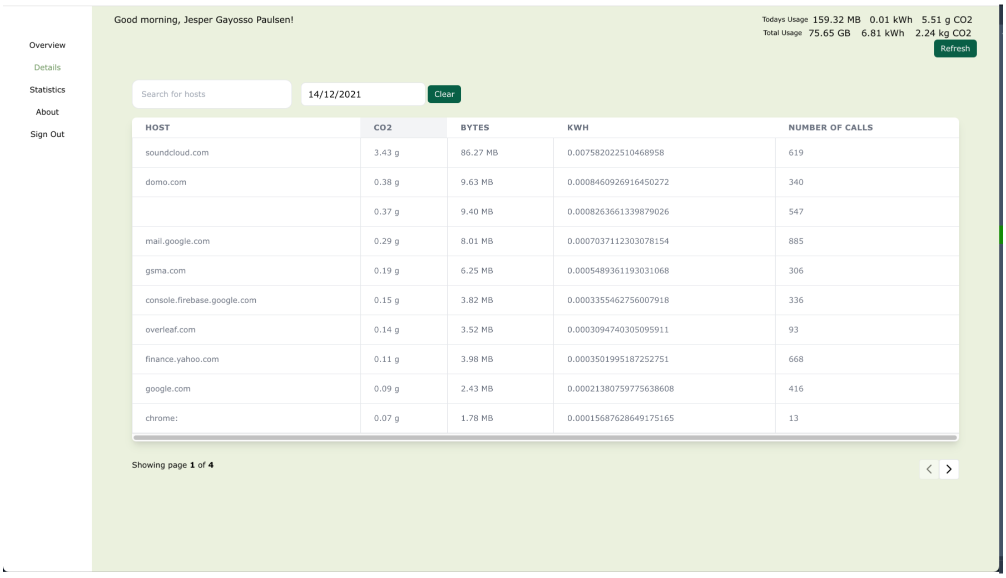

Figure 10.23: Final implementation of the DataVisualizer's details table. Here the results are sorted based on CO_2e emissions the 14th of December 2021.

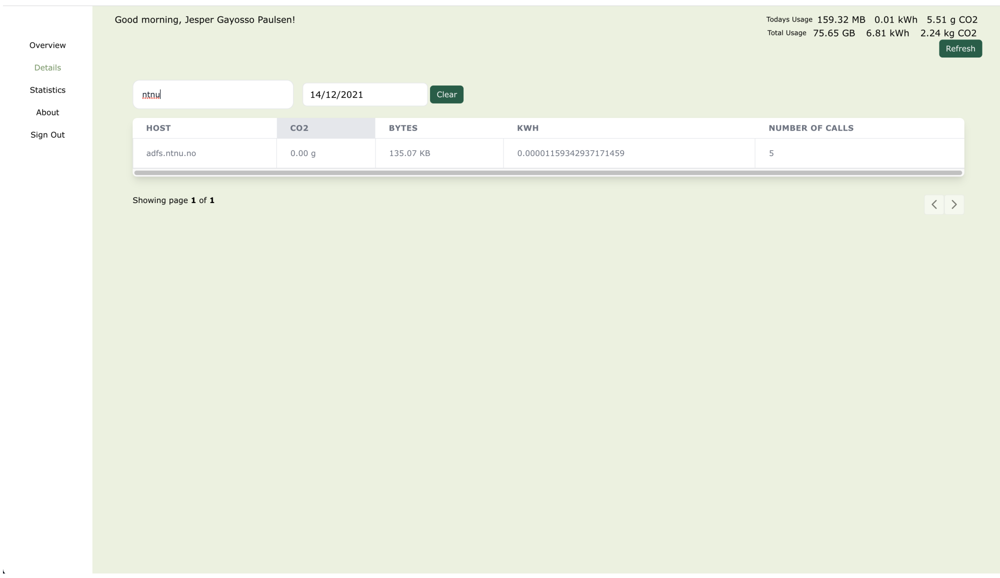

Figure 10.24: Final implementation of the DataVisualizer's details table. Here the results matching the input string of *ntnu* is shown.

10.10.4 Statistics

The final page of the DataVisualizer is the statistics page. This page is shown in Figure 10.25.

Chart The Chart in Figure 10.25 shows the user's usage compared to the average usage of the other users per day.

Dropdown The dropdown on the top lets the user select either CO_2, size, kWh or number of requests. In Figure 10.26 *number of requests* is selected. Compared to the statistics in Figure 10.25, this example shows how fewer network requests sent on *Sunday* where responsible for significantly more CO_2 pollution, compared to other days.

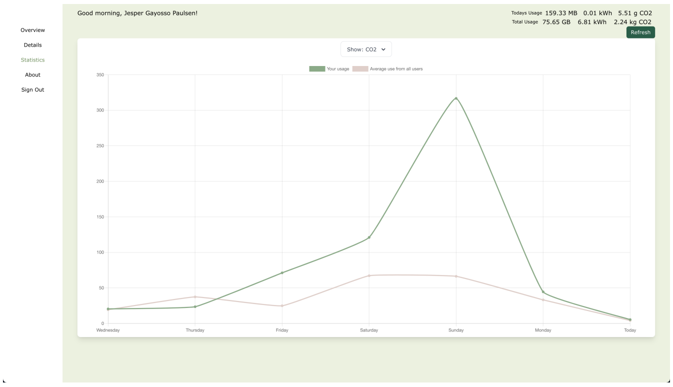

Figure 10.25: Final implementation of the DataVisualizer's statistics page.

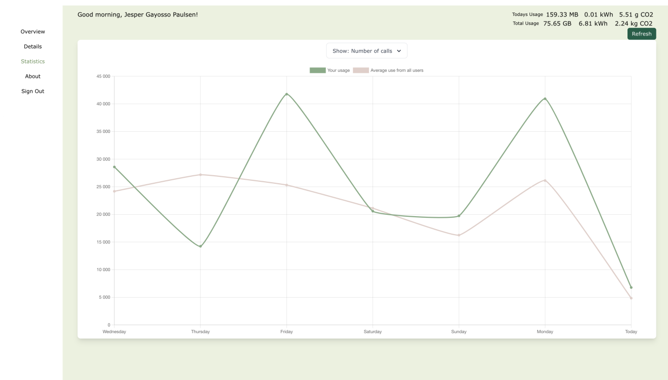

Figure 10.26: Final implementation of the DataVisualizer's statistics page. Here the user have selected *Number of calls* in the dropdown menu.

10.11 Project's Website and Content Management System

A website is created to inform about the project and let the participants sign up. To quickly change the website's content, a Content Management System (CMS) was hooked up with the website. A CMS makes it easy to write text and create content on a webpage. Thus, making it possible to avoid hard-coding the content in the actual code. As the website and CMS is not directly connected to the success of this experiment, the details of the implementation will be presented in Appendix 5.

10.12 Summary

This chapter has presented the system's architecture, the technologies used, and the design phase. Finally, the implementation of the DataPresenter and the DataVisualizer has been presented in detail.

11 Validation of the Functional and the Non-Functional Requirements

This chapter outlines the testing and validation of the functional and the non-functional requirements. Thus, confirming that the final solution meets the criteria set by the requirements.

11.1 Approach

Before creating the system, the functional requirements were divided into smaller issues. Dividing tasks into smaller tasks makes it easier to follow an agile development, using the *INVEST (Independent, Negotiable, Valuable, Estimable, Small, Testable)* acronym defined by extreme programming to split tasks [144]. In order to manage the tasks, a kanban board was used. A kanban board aids in the development phase by visualizing the different tasks, and their status [121]. As I was the only developer, a mono branch approach was chosen without pull requests. When there is only one developer, pull requests make the developing process more cumbersome than the value it produces [19].

11.2 Validation of the Functional Requirements

When dividing the requirements into smaller tasks, each task was stored with a reference to their functional requirement. This made it possible to monitor the completion of the different functional requirements. Thus, each requirement was tested after completing the tasks belonging to it. New tasks were created if the testing did not satisfy the functional requirement. These tasks were then added to the backlog and prioritized based on the priority attribute of the functional requirement. The validation of the functional requirements is presented in Table 11.1.

ID	Functional Requirement	Status	Comment
FR 1	The system should detect amount of data being sent	Completed	The DataCollector is able to detect the amount of data being sent.
FR 2	The system should detect the location the data is being sent to	Completed	The DataCollector + DataAnalyzer is able to detect the location of the data.
FR 3	The system should require authentication of the user	Completed	Firebase Authentication is used, and tests are written to ensure people only have access to their own data.
FR 4	The system should make it possible for the user to select what data that is being logged	Partially	It is possible for the user to blacklist domains that the DataCollector should not log. But it is not possible to select what kind of information to log from different websites.
FR 5	The system should update in real time	Completed	Cloud Firestore ensures that data is updated in real time.
FR 6	The system should make it possible for the user to see what action caused what pollution	Completed	By the use of the real-time updates, the badge, push notifications, the world map and the details table, the user should be able to understand what action that caused what pollution.
FR 7	The system should make it possible for the user to explore its pollution caused around the world	Completed	By the use of the world map the user is able to explore its data
FR 8	The system should make it possible for users to compare pollution to the average of other users	Completed	By the use of the Statistics Chart in the DataVisualizer, and the comparisons in the DataPresenter, the user is able to compare its own usage to the others.
FR 9	The system should generate reports for the user, based on its usage	Completed	Daily reports are generated by the DataCollector and shown in the DataPresenter
FR 10	The system should provide the user with tips on how to reduce the pollution	Completed	Tips are fetched from the CMS and presented in the DataPresenter
FR 11	The system should send push notifications with usage updates	Completed	Push notifications are sent from the DataCollector.

Table 11.1: Validation of the Functional Requirements for the system

11.3 Validation of Non-Functional Requirements

The validation of the non-functional requirements was executed at the end of the project. The status of the non-functional requirements is presented in Table 11.2. In total, four out of seven requirements were completed. Two requirements were partially completed, and one was not completed. In the comments column, a comment about the status of each requirement is given.

ID	**Non-Functional Requirement**	**Status**	**Comment**
NFR 1	Implementation time less than 3 months	Completed	The system was implemented in less than 3 months. Some bug fixes was done outside of the scope of the months, but the requirement is still seen as completed.
NFR 2	The system should be secure	Completed	By the use of Firebase Authentication, Firestore Security Rules, Chrome's identity API and tests the system is said to be secure.
NFR 3	The system should be runable on Windows, Mac OS and Linux	Completed	As the applications are runnable in Chrome, this requirement is completed.
NFR 4	The system should be easy to install	Not Completed	It is not easy to install and set up an application trough the chrome web store. This was outside my control.
NFR 5	The system should report the data without manual intervention from the user	Completed	The system reports the data automatically after it has been installed.
NFR 6	The system should be easy to test	Partially	The system as a whole is not easy to test, but the API is testable, which was thought to be the most important part to test.
NFR 7	The system should have a low carbon footprint	Partially	Decisions have actively been taken to make the system as climate friendly as possible without limiting the functionality of the system within the given timeframe, but there is believed to be more parts of the architecture that can be optimized.

Table 11.2: Validation of the Non-Functional Requirements of the system

11.4 Summary

In this chapter, the status of the functional and the non-functional requirements have been presented. Of the 11 functional requirements, ten were completed. One functional requirement is marked as *Partially Completed*, as more functionality could have been added to satisfy the requirement. On the other hand, only four out of the seven non-functional requirements were completed. One non-functional requirement is marked as *Not Completed*, and two are marked as *Partially Completed*.

12 Execution

This chapter will explain the two data generation methods. First, the two questionnaires will be presented. Then, the experiment using the Chrome extension will be explored.

12.1 Norwegian Centre for Research Data (NSD)

This thesis had to be reported to NSD as one of the data generation methods collected and processed personal data. The form that was sent to and approved by NSD included information about what personal data would be collected, the duration of the project, how the data would be stored and processed, the reason for collecting the data, and the project's goal.

12.2 Data Generation Method 1: Questionnaires

The first method used as a data generation in this project was questionnaires. In total, two questionnaires were created. The questions of the questionnaires are presented in Appendix 7. The questionnaires were distributed through Nettskjema[1], as self-administrated questionnaires [109]. The questionnaires were chosen as one of the data generation methods as they make it easy to collect standardized data from a large number of people, reducing the time spent on data preparation [109].

12.2.1 Response Types

Different response types were used in the Questionnaires. As suggested by Oates, the questionnaires were designed to limit the type of response types [109].

Yes/No Some questions were yes/no questions. These questions were used to control the visibility of other questions. They were only used on the first questionnaire.

Drop-down The drop-down response type was chosen for questions containing many possible alternatives. These were chosen because they reduce the visual size of the questionnaires.

[1] `https://nettskjema.no/`

Radio Buttons and Checkboxes Radio buttons were chosen as an alternative to drop-down lists when the answers contained seven elements or less. This reduced the number of clicks required to answer a given question and provided the participant with an overview of the possible answers. One of the questions required checkboxes instead of radio buttons, as the question demanded multiple answers to be selected.

Open Questions Closed questions were chosen for the majority of the questions to limit the effort required by the participants [109]. However, one open question was chosen for the last question in the second questionnaire. This question allowed the participant to provide feedback to the project, aiding in the qualitative analysis of the project.

Likert Scale In order to capture the degree of agreement or disagreement, the Likert scale was used [107, 109]. The Likert scale is a psychometric scale and is "one of the most fundamental and frequently used psychometric tools in educational and social sciences research" [76]. One of the greatest benefits of using the Likert scale is that the participants can relate to similar experiences when answering the questions. Therefore, the Likert scale was used as a response type for most of the questions.

12.2.2 First Questionnaire

The first questionnaire examined the users' Internet behavior, knowledge of the Internet, digital climate footprint, and willingness to change. It consisted of seven parts, and in total, 41 questions.

12.2.3 Second Questionnaire

The second questionnaire aimed to see if the participants had changed any habits or gained new knowledge or experience during the project. It consisted of six parts and 34 questions. The questions focused on the participants' Internet usage and the different features of the plugin. Part five of the questionnaire was identical to part five of questionnaire one. Thus, making it possible to compare the results of the two questionnaires.

12.3 Data Generation Method 2: Chrome extension

This thesis's second data generation method was to collect the users' digital carbon footprint using the system developed. During the first week of the experiment, the participants did not receive feedback on their Internet usage, as shown in Figure 12.1. In the background, the plugin reported the amount of data the participants were using and the number of seconds they were active per day. During the second week, the participants were given feedback on their usage as shown in Figure 12.2. By not giving feedback to the users during the first phase, the data collected from the second week could be compared to the data from the first week to look for a change in usage trends.

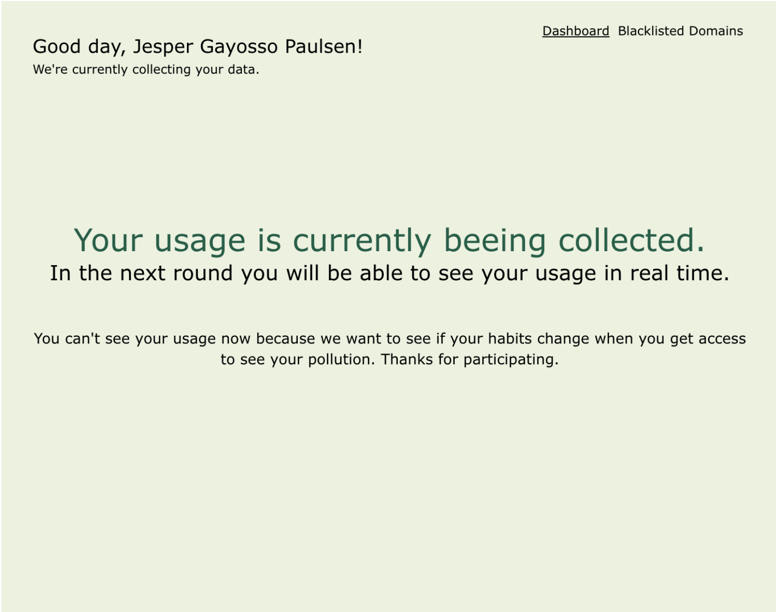

Figure 12.1: The information shown to the user during the first week of the experiment.

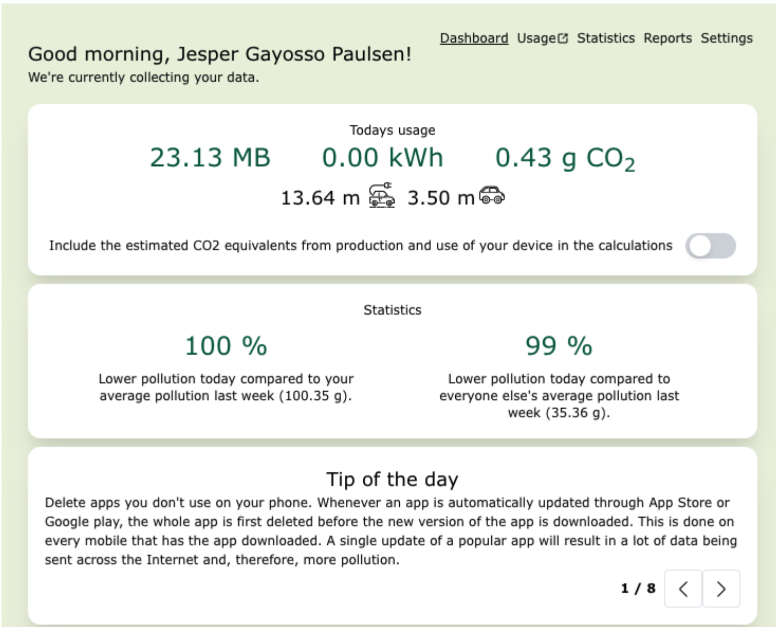

Figure 12.2: The information shown to the user during the second week of the experiment.

12.4 Data Analyzes

At the end of the project, the data from the data generation methods were analyzed. The results from the questionnaires were examined to understand if the users had become more aware of their digital carbon footprint. The results from the extension were studied to see if the users changed their digital behavior. The results will be presented in Chapter 13.

12.5 Summary

In this Chapter, the execution of the project has been explained in detail. The project used two data generation methods; Questionnaires and the System developed. The questionnaires were used to examine if the user intentionally changed their behavior, while the system developed was used to look for an actual change in their digital behavior.

13 Results

This chapter presents the results from the two data generation methods presented in Chapter 12. First, the distribution of the participants that were part of the project will be presented. Second, the results from the questionnaires will be displayed. Last, the results from the data collected through the extension will be explored.

13.1 Participants

There were in total 72 people that signed up for the project. Of these, 43 participants answered the first questionnaire. From the 43 participants in the first questionnaire, 27 people installed the extension. The main reason for people not installing the extension was that they did not use Chrome, had problems installing the extension, or were not allowed by their company to install software on their work computers. The final questionnaire was only sent to the people who used the extension. In total 20 participants answered the final questionnaire.

13.1.1 Distribution of Participants Answering the First Questionnaire

The distribution of the participants answering the first questionnaire can be seen in Figure 13.1, Figure 13.2, Figure 13.3 and Figure 13.4. The figures show that *males*, *females*, and *others* answered the first questionnaire. Most of the participants were between 25-44 years old and from Norway. The figures also show that most of the participants have experience developing or designing websites and apps that use the Internet.

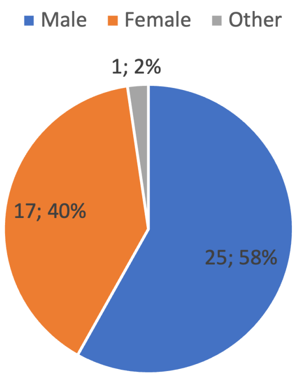

Figure 13.1: Gender of participants answering the first questionnaire

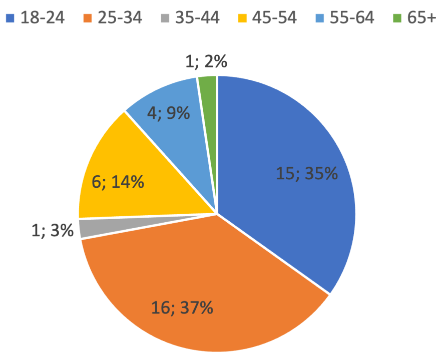

Figure 13.2: Age of participants answering the first questionnaire

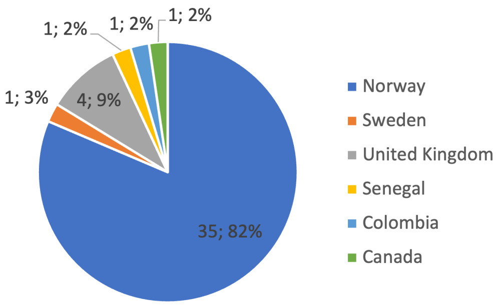

Figure 13.3: Participants' country of origin

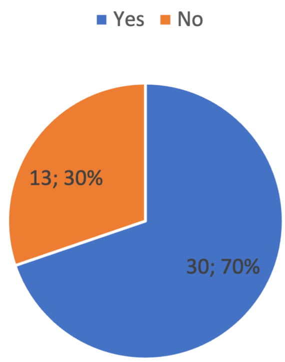

Figure 13.4: Participants' that have studied, or are studying, and/or have developed or designed websites and/or apps that use the Internet

13.1.2 Distribution of Participants Answering the Second Questionnaire

Two did not answer the first questionnaire of the 20 people answering the final questionnaire. These two answers were therefore excluded from the results. The distribution of the participants answering the second questionnaire is shown in Figure 13.1, Figure 13.2, Figure 13.3 and Figure 13.4. The figures show that 67% of the participants answering the final questionnaire were males. It also reveals that most participants were between 25-34 years old and from Norway. The majority of the second questionnaire participants have experience developing or designing websites and apps that use the Internet.

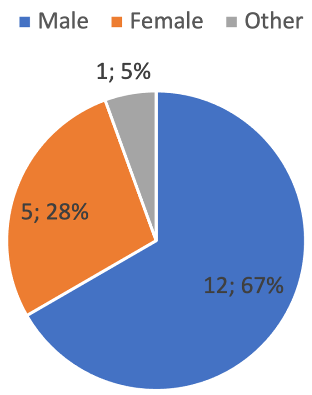

Figure 13.5: Gender of participants answering the second questionnaire

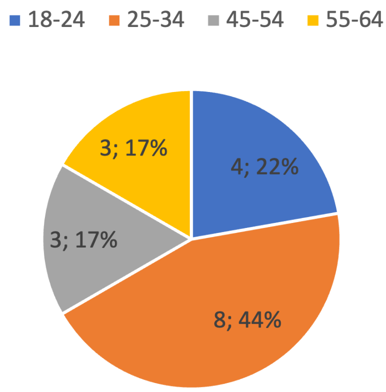

Figure 13.6: Age of participants answering the second questionnaire

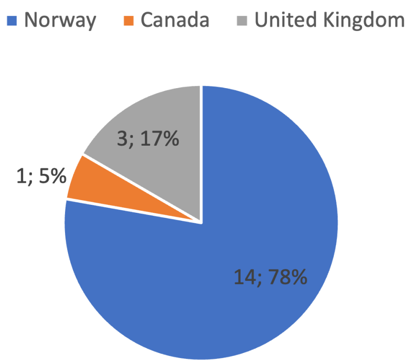

Figure 13.7: Participants' country of origin

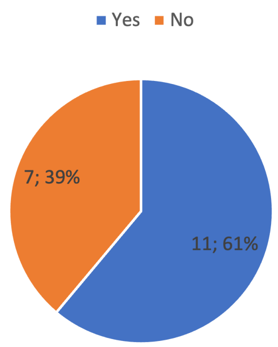

Figure 13.8: Participants' that have studied, or are studying, and/or have developed or designed websites and/or apps that use the Internet

13.2 Results From the Questionnaires

In this section, the results from the questionnaires are presented. The tables present either the results from a single questionnaire or from both questionnaires compared. When analyzing both questionnaires compared, Wilcoxon Signed Rank Test will be used to look for significant results [151]. The Wilcoxon Signed Rank test is non-parametric, used to analyze paired data. The test was chosen because it is influential on data that is not normally distributed, and it works with ordinal data [138].

In addition to analyzing questions across the questionnaires, the participants were divided into groups. The groups were defined as follows:

- Age: Split into two groups, below and above 35.
- Gender: Split into two groups: Male and Female/Other. As there were more males than females, Other was included in the Female group.
- Daily time spent online: Divided into two groups based on their aggregated answers from part 2 of Questionnaire 1. The distribution is shown in Figure 13.9. The groups were split on the median of the results. For questionnaire one, the median was 11.5, and for questionnaire two, the median was 11.25.
- Experience with developing or designing or studying websites and or apps that use the Internet: Divided into two groups based on whether or not the participant answered yes on question 1.5 and 1.6 in Questionnaire 1.
- Willingness to change behavior on the Internet to save the Environment: Divided based on the willingness to change. Participants who answered agreeing to the *Question 6.1: I'm willing to change my behavior on the Internet to save the environment* were in one group, and the participants who answered neutral or disagreeing to this question were in the other group.
- For questionnaire 2; Estimated CO2 emissions during the project. The groups were divided on the median based on the estimated CO2 emissions caused by their data usage.

The Mann-Whitney U test is used when analyzing the questions based on the groups presented. The reason for using this test is that it is non-parametric, works with ordinal data, and does not require the two independent distributions to be normally distributed [92].

When analyzing the results from the Likert scale questions, the five-point scale will be used. However, to increase the readability of the results, the results will be presented with a three-point Likert scale. For these experiments, a significance level of 0.05 was chosen. The experiments presented will only include the results that were either significant or borderline to significant. The significant results will be highlighted in bold. When presenting the Likert scale questions, the following definitions will be used:

- **n**: Number of respondents to a given question. If the results are grouped, it will represent the number of participants in a group. Otherwise, it will represent the total amount of answers received for the question.
- **D**: Number of Disagreeing participants. This will be "Strongly Disagree" and "Disagree" summed.
- **N**: Number of Neutral participants.
- **A**: Number of Agreeing participants. This will be "Agree" and "Strongly Agree" summed.
- **p**: Probability value

13.2.1 Results from Questionnaire 1

The results from the first questionnaire will now be explored.

Part 2: Information About Your Internet Usage

The results from Part 2 are outlined in the histogram shown in Figure 13.9. The Figure summarizes the answers to the questions shown in Table 13.1. When calculating the total usage, each value from the results of each question is converted into a numeric value, using the conversion values shown in Table 13.2. From the histogram, it can be derived that most of the participants use the services listed in Table 13.1 for a total of 13 hours or less per day. None of the participants uses the services for less than 5 hours per day, and one participant uses the services for 29 hours per day. This number is greater than the 24 hours of a day because some of the services can be used simultaneously.

ID	Question
2.1	On a daily basis, how many hours do you spend streaming music on platforms like Spotify and Soundcloud?
2.2	On a daily basis, how many hours do you spend on platforms like TikTok and YouTube?
2.3	On a daily basis, how many hours do you spend on platforms like Facebook, LinkedIn, and Twitter?
2.4	On a daily basis, how many hours do you spend on streaming platforms like Netflix and HBO Max?
2.5	On a daily basis, how many hours do you spend in digital meetings on platforms like Zoom and Teams?
2.6	On a daily basis, how many hours do you spend on regular browsing?
2.7	On a daily basis, how many hours do you spend on online gaming?

Table 13.1: The questions used to create the Histogram shown in Figure 13.9

.

Original value	**Converted value**
Not using it	0
Less than 1 hour	0.5
1-2 hours	2
3-4 hours	4
5-6 hours	6
7-8 hours	8
9-10 hours	10
11-12 hours	12
13-14 hours	14
More than 14 hours	15

Table 13.2: The values used to convert to the answers from the questions in part 2 to numeric values.

.

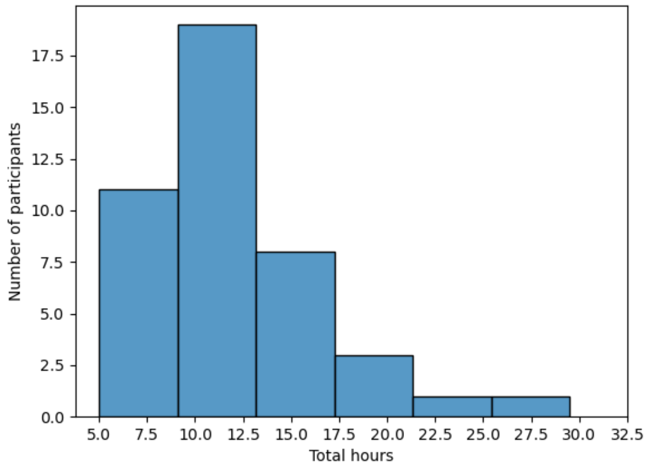

Figure 13.9: Distribution of the participants Internet usage.

Part 3: Background Questions

The results from part 3 of the first questionnaire are presented in Figure 13.10 and Figure 13.11. Interestingly, 95% of the participants answered that they knew that the Internet was polluting before entering the study. Furthermore, 54% assumed that the Internet is causing the same pollution as the aviation industry. Ten respondents reckoned that the aviation industry was responsible for less than half of the aviation industry, and ten participants assumed the Internet was responsible for more than twice as much as the aviation industry. Based on the findings presented in Chapter 5, there is not an exact answer to this question. However, the answer "Roughly the same as the aviation industry" seems to be the most fitting based on the research presented in Chapter 5.

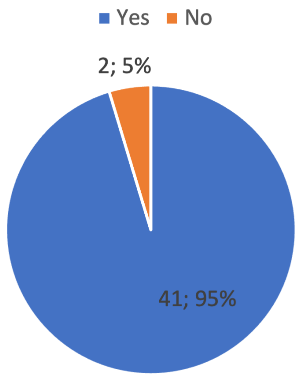

Figure 13.10: Results from Question 3.1: Before entering this study - did you know that the Internet pollutes?

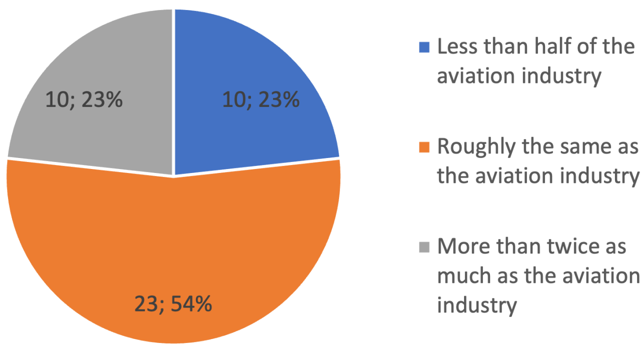

Figure 13.11: Results from Question 3.2: Compared to the aviation industry, how much pollution do you think that Internet usage causes?

Part 4: The Internet and the Environment

The results from part 4 is presented in Table 13.3. The results show that most of the participants care about the environment. It also shows that 65% of the participants know how the cloud works. Furthermore, most of the respondents answered that it is important for them to make environmentally friendly choices and that they want to learn how to reduce their digital carbon footprint. All respondents answered that they wanted to learn how the Internet pollutes, except one participant. Also, 41 of the 43 participants wanted to learn how they could reduce their digital carbon footprint.

ID	Question	n	D	N	A
4.1	I care about the environment	43	0 (0%)	1 (2.32%)	42 (97.67%)
4.2	I know how the "cloud" works	43	8 (18.60%)	7 (16.27%)	28 (65.11%)
4.3	It is important for me to make environmental friendly choices	43	0 (0%)	4 (9.30%)	39 (90.69%)
4.4	I know how I can reduce my digital carbon footprint.	43	27 (62.79%)	9 (20.93%)	7 (16.27%)
4.5	I want to learn how I can reduce my digital carbon footprint.	43	0 (0%)	2 (4.65%)	41 (95.35%)
4.6	I want to learn how the Internet pollutes	43	0 (0%)	1 (2.32%)	42 (97.67%)

Table 13.3: Results from part 4: The Internet and the Environment

Part 5: The Environment

In Table 13.4, the results from questionnaire 1 part 5 is shown. The results show that roughly half of the participants feel guilty when flying, while none feel guilty when using the Internet. The results also show that most respondents do not think about the environment when performing different actions on the Internet. However, one participant thinks about the environment before uploading a file to the cloud.

ID	Question	n	D	N	A
5.1	I feel guilty about flying	43	11 (25,58%)	8 (18,60%)	24 (55,81%)
5.2	I feel guilty about using the Internet.	43	39 (90,70%)	4 (9,30%)	0 (0%)
5.3	I think about the environment before I visit a website	43	43 (100%)	0 (0%)	0 (0%)
5.4	I think about the environment before I post an image to social media	43	43 (100%)	0 (0%)	0 (0%)
5.5	I think about the environment before I watch a video on the Internet	43	42 (97,67%)	1 (2,33%)	0 (0%)
5.6	I think about the environment before I stream a video	43	41 (95,34%)	2 (4,65%)	0 (0%)
5.7	I think about the environment before I upload a file to the cloud	43	42 (97,67%)	0 (0%)	1 (2,33%)

Table 13.4: Results from part 5: The Environment

Part 6: Willingness to Change

The results from the 6th part of the first questionnaire are shown in Table 13.5. More than half of the participants are willing to change their behavior on the Internet and reduce the amount of data they use to save the environment. The majority of the respondents want to learn where in the world they pollute when using the Internet. In contrast, less than half of the participants want to stop visiting web pages with many images and videos. 72% of the participants are willing to stop visiting websites with a dirty carbon footprint, given that they get information about the websites' carbon footprint. Of the questions in part 6, the statement the participants least agree with is "*I'm willing to reduce the time I spend on the Internet if it reduces my digital carbon footprint.*"

ID	Question	n	D	N	A
6.1	I'm willing to change my behavior on the Internet to save the environment	43	7 (16,28%)	11 (25,58%)	25 (58,14%)
6.2	I'm willing to reduce the amount of data I'm using on the Internet if it will reduce my digital carbon footprint	43	4 (9,3%)	14 (32,56%)	25 (58,14%)
6.3	I'm willing to reduce the time I spend on the Internet if it reduces my digital carbon footprint.	43	11 (25,58%)	15 (34,88%)	17 (39,53%)
6.4	I'm willing to have less stored in the cloud if it will reduce my digital carbon footprint	43	8 (18,60%)	11 (25,58%)	24 (55,13%)
6.5	I want to learn where in the world I pollute when using the Internet	43	1 (2,32%)	6 (13,95%)	36 (83,72%)
6.6	I'm willing to stop visiting websites that pollute much if I get feedback on how much they pollute	43	3 (6,98%)	9 (20,93%)	31 (72,09%)
6.7	I feel images, videos and animations are important to make websites more interesting	43	5 (11,63%)	17 (39,53%)	21 (48,84%)
6.8	I'm willing to visit webpages that looks less "fancy" and uses less images and videos if it will reduce my digital carbon footprint	43	7 (16,28%)	17 (39,53%)	19 (44,19%)
6.9	I wish search engines like Google sorted the relevant search results based on how much they pollute	43	8 (18,60%)	9 (20,93%)	26 (60,46%)

6.10	I wish cloud providers made it easier to delete images, videos, and documents.	43	5 (11,63%)	10 (23,26%)	28 (65,12%)

Table 13.5: Results from part 6: Willingness to change

Part 7: Questions for Developers and Designers

In the final part of the first questionnaire, the number of participants (n) is slightly lower than the other questions. These questions were only shown to the participants who have developed or designed websites or apps that use the Internet and participants who have studied how to develop or design websites or apps that use the Internet. Of the 23 participants in this group, only one agreed with the statement *"While studying, I learned about how the Internet pollutes."* None of the participants learned how to create websites or apps that pollute less. However, two of them think of how to create websites or apps that pollute less.

ID	Question	n	D	N	A
7.1	While studying, I learned about how the Internet pollutes.	23	21 (91,3%)	1 (4,35%)	1 (4,35%)
7.2	While studying, I learned how I could create websites and apps that pollute less.	23	23 (100%)	0 (0%)	0 (0%)
7.3	When creating or designing websites and/or apps, I think about how to create websites and/or apps that pollute less.	28	22 (78,57%)	4 (14,29%)	2 (7,14%)

Table 13.6: Results from part 7: Questions for developers and designers

Age

Table 13.7 shows the questions from Questionnaire 1, grouped by age. The results show that the older group thinks it is important to make environmentally friendly choices, while the younger group is neutral. Almost all the participants want to learn how the Internet pollutes, except for one participant in the younger group. A more significant percentage of the younger group does not know how to reduce their digital carbon footprint compared to the older ones. The older group is also more interested in reducing their digital carbon footprint. In comparison, 19% of the younger group do not want to change their behavior on the Internet to save the environment.

ID	Question	Group	n	D	N	A	p
4.3	It is important for me to make environmental friendly choices	< 35	31	0 (0.0%)	4 (12.9%)	27 (87.1%)	0.07729
		>= 35	12	0 (0.0%)	0 (0.0%)	12 (100.0%)	
4.4	I want to learn how the Internet pollutes	< 35	31	0 (0.0%)	1 (3.23%)	30 (96.77%)	0.07678
		>= 35	12	0 (0.0%)	0 (0.0%)	12 (100.0%)	
4.5	I know how I can reduce my digital carbon footprint	< 35	31	22 (70.97%)	5 (16.13%)	4 (12.9%)	0.05691
		>= 35	12	5 (41.67%)	4 (33.33%)	3 (25.0%)	
4.6	I want to learn how I can reduce my digital carbon footprint.	< 35	31	0 (0.0%)	2 (6.45%)	29 (93.55%)	**0.01884**
		>= 35	12	0 (0.0%)	0 (0.0%)	12 (100.0%)	
6.1	I'm willing to change my behavior on the Internet to save the environment	< 35	31	6 (19.35%)	9 (29.03%)	16 (51.61%)	0.07503
		>= 35	12	1 (8.33%)	2 (16.67%)	9 (75.0%)	
6.10	I wish cloud providers made it easier to delete images, videos, and documents.	< 35	31	5 (16.13%)	8 (25.81%)	18 (58.06%)	0.08688
		>= 35	12	0 (0.0%)	2 (16.67%)	10 (83.33%)	

Table 13.7: Results from the first questionnaire, grouped by age. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

Gender

The significant results when dividing the participants by gender are shown in Table 13.8. To make two groups and even out the sample size, other was added to the female group. The results show that the males care more about the environment and know how the cloud works more accurately than the females. Also, the females are more unsure about how to reduce their digital carbon footprint compared to the males.

ID	Question	Group	n	D	N	A	p
4.1	I care about the environment	Male	25	0 (0.0%)	0 (0.0%)	25 (100.0%)	**0.01447**
		Female + Other	18	0 (0.0%)	1 (5.56%)	17 (94.44%)	
4.2	I know how the "cloud" works	Male	25	2 (8.0%)	1 (4.0%)	22 (88.0%)	**0.00142**
		Female + Other	18	6 (33.33%)	6 (33.33%)	6 (33.33%)	
4.5	I know how I can reduce my digital carbon footprint	Male	25	12 (48.0%)	7 (28.0%)	6 (24.0%)	**0.01537**
		Female + Other	18	15 (83.33%)	2 (11.11%)	1 (5.56%)	

Table 13.8: Results from the first questionnaire, grouped by gender. Showing questions where the probability value (p) is significant (p < 0.5, highlighted in bold) and borderline to significant values (p < 0.1).

Daily Time Spent Online

The results when dividing participants based on their answers on aggregated time spent online are shown in Table 13.9, based on the results presented in Section 13.2.1. The participants with an aggregated time lower than the median of 11.5 are in one group, and the rest is in the other group. The results show that more of the persons with a high total time know how the cloud works, compared to those with a lower time. It also reveals that none of the participants, except from one of the participants with a high total time spent on the Internet, disagrees with the statement *"I think about the environment before I watch a video on the Internet"*. Compared to the other group, the participants with a lower total time are also more positive to have less stored in the cloud if it will reduce their digital carbon footprint.

Experience with developing/designing websites

Table 13.10 shows the results from dividing the group based on their experience with developing or designing websites or apps that use the Internet. The results show that

ID	Question	Group	n	D	N	A	p
4.2	I know how the "cloud" works	< 11.5	21	6 (28.57%)	4 (19.05%)	11 (52.38%)	**0.04224**
		>= 11.5	22	2 (9.09%)	3 (13.64%)	17 (77.27%)	
5.5	I think about the environment before I watch a video on the Internet	< 11.5	21	21 (100.0%)	0 (0.0%)	0 (0.0%)	0.08309
		>= 11.5	22	21 (95.45%)	1 (4.55%)	0 (0.0%)	
6.4	I'm willing to have less stored in the cloud if it will reduce my digital carbon footprint	< 11.5	21	3 (14.29%)	4 (19.05%)	14 (66.67%)	0.08218
		>= 11.5	22	5 (22.73%)	7 (31.82%)	10 (45.45%)	

Table 13.9: Results from the first questionnaire, grouped by time spent on the Internet. Showing questions where the probability value (p) is significant (p < 0.5, highlighted in bold) and borderline to significant values (p < 0.1).

more of the participants with experience know how the cloud works than those with no experience.

ID	Question	Group	n	D	N	A	p
4.2	I know how the "cloud" works	None	13	5 (38.46%)	3 (23.08%)	5 (38.46%)	**0.02705**
		Experi.	30	3 (10.0%)	4 (13.33%)	23 (76.67%)	

Table 13.10: Result from the first questionnaire, grouped by their experience with developing and/or designing websites and/or apps that use the Internet.

Willingness to Change

When dividing the participants by their answers to *Question 6.1: I'm willing to change my behavior on the Internet to save the environment*, we get the results presented in Table 13.11. Participants that answered they agreed to the question are placed in the "High willingness" group, and the rest is placed in the "Low willingness" group.

The results show that people with high willingness think it is more important to make environmentally friendly choices than the other group. Furthermore, they are also more interested in learning how the Internet pollutes and how they can reduce their digital carbon footprint, compared to the other group. 38% of the participants in the group with low willingness do not feel guilty when flying, while 64% of the participants in the high willingness group answered that they do. The high willingness group is more willing to reduce the time and data they spend on the Internet than the other group.

In addition, the respondents with high willingness are also more interested in learning where in the world they pollute when they use the Internet, compared to the group with low willingness. Most of the participants with a high willingness to change behavior think that search providers should sort relevant search results based on how much they pollute. The exact number of respondents in this group also thinks that cloud providers should simplify deleting files from the cloud. On the other hand, six participants with low willingness do not think search results should be sorted based on how much they pollute, and four respondents in the same group do not think that the cloud providers should make it easier to delete files.

ID	Question	Group	n	D	N	A	p
4.3	It is important for me to make environmental friendly choices	Low	18	0 (0.0%)	4 (22.22%)	14 (77.78%)	**0.01247**
		High	25	0 (0.0%)	0 (0.0%)	25 (100.0%)	
4.4	I want to learn how the Internet pollutes	Low	18	0 (0.0%)	1 (5.56%)	17 (94.44%)	**0.00605**
		High	25	0 (0.0%)	0 (0.0%)	25 (100.0%)	
4.5	I want to learn how I can reduce my digital carbon footprint.	Low	18	0 (0.0%)	2 (11.11%)	16 (88.89%)	0.07253
		High	25	0 (0.0%)	0 (0.0%)	25 (100.0%)	
5.1	I feel guilty about flying.	Low	18	7 (38.89%)	3 (16.67%)	8 (44.44%)	0.09091
		High	25	4 (16.0%)	5 (20.0%)	16 (64.0%)	
5.6	I think about the environment before I stream a video	Low	18	18 (100.0%)	0 (0.0%)	0 (0.0%)	0.07857
		High	25	23 (92.0%)	2 (8.0%)	0 (0.0%)	
6.2	I'm willing to reduce the amount of data I'm using on the Internet if it will reduce my digital carbon	Low	18	3 (16.67%)	11 (61.11%)	4 (22.22%)	**0.00004**
		High	25	1 (4.0%)	3 (12.0%)	21 (84.0%)	
6.3	I'm willing to reduce the time I spend on the Internet if it reduces my digital carbon footprint.	Low	18	7 (38.89%)	9 (50.0%)	2 (11.11%)	**0.0034**
		High	25	4 (16.0%)	6 (24.0%)	15 (60.0%)	
6.4	I'm willing to have less stored in the cloud if it will reduce my digital carbon footprint	Low	18	5 (27.78%)	6 (33.33%)	7 (38.89%)	0.05239
		High	25	3 (12.0%)	5 (20.0%)	17 (68.0%)	

6.5	I want to learn where in the world I pollute when using the Internet	Low	18	1 (5.56%)	4 (22.22%)	13 (72.22%)	**0.00544**
		High	25	0 (0.0%)	2 (8.0%)	23 (92.0%)	
6.6	I'm willing to stop visiting websites that pollute much if I get feedback on how much they pollute	Low	18	3 (16.67%)	7 (38.89%)	8 (44.44%)	**0**
		High	25	0 (0.0%)	2 (8.0%)	23 (92.0%)	
6.9	I wish search engines like Google sorted the relevant search results based on how much they pollute	Low	18	6 (33.33%)	5 (27.78%)	7 (38.89%)	**0.00116**
		High	25	2 (8.0%)	4 (16.0%)	19 (76.0%)	
6.10	I wish cloud providers made it easier to delete images, videos, and documents.	Low	18	4 (22.22%)	5 (27.78%)	9 (50.0%)	**0.04575**
		High	25	1 (4.0%)	5 (20.0%)	19 (76.0%)	

Table 13.11: Results from the first questionnaire, grouped by their willingness to change their behaviour to save the environment. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

13.2.2 Results from Questionnaire 2

In this section, the results from the second questionnaire will be presented.

Part 2: Information about your Internet usage

The results from Part 2 of Questionnaire 2 are presented in Table 13.12. The number of participants per answer varies as participants who had not used the mentioned platform in a given question were told not to answer the question. In total, 7 of the participants answered agreeing to one or more of the questions 2.1 to 2.7. The results show that most of the participants have not reduced the time they spend streaming music. Furthermore, half of the respondents disagree with the statement *"After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom and Teams"*. However, some of the participants have reduced the time they spend on platforms like YouTube and LinkedIn. At the same time, 72% say that they think about how they can reduce their digital carbon footprint when using the Internet.

ID	Question	n	D	N	A
2.1	After joining the project, I spend less time streaming music on platforms like Spotify and Soundcloud	16	11 (68.75%)	5 (31.25%)	0 (0.0%)
2.2	After joining the project, I spend less time on platforms like TikTok and YouTube	13	8 (61.53%)	3 (23.08%)	2 (15.38%)
2.3	After joining the project, I spend less time on platforms like Facebook, LinkedIn, and Twitter	17	9 (52.94%)	4 (23.53%)	4 (23.53%)
2.4	After joining the project, I spend less time on streaming platforms like Netflix and HBO Max	16	8 (50.0%)	4 (25.0%)	4 (25.0%)
2.5	After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom and Teams	17	9 (52.94%)	5 (29.41%)	3 (17.65%)
2.6	After joining the project, I spend less time on regular browsing	18	5 (27.78%)	9 (50.0%)	4 (22.22%)
2.7	After joining the project, I spend less time on online gaming	12	7 (58.33%)	5 (41.67%)	0 (0.0%)
2.8	After joining the project, I think about how I can reduce my digital carbon footprint when using the Internet	18	1 (5.56%)	4 (22.22%)	13 (72.22%)

Table 13.12: Results from part 2: Information about your Internet usage

Part 3: Plugin Features

In Table 13.13, the results from Part 3 of Questionnaire 2 is presented. It shows that most of the participants found the different plugin features useful; however, push notifications were the least popular feature. Overall, the two features that scored best were the world map functionality and the chart that compared the user's usage to everyone else's usage. When looking at the results, the different features of the system can be grouped based on their popularity. This will aid in understanding what features that should be prioritized when creating similar systems in the future. In order to calculate the popularity of the features, the total sum from the Likert Scale questions is summed up. The result is shown in Table 13.14. Overall, the scores for the different features were quite even, except for the push notification feature.

ID	Question	n	D	N	A
3.1	It was useful to recieve tips on how I can reduce my digital carbon footprint	18	1 (5.56%)	4 (22.22%)	13 (72.22%)
3.2	It was useful to see my usage and pollution caused in real-time	18	2 (11.11%)	4 (22.22%)	12 (66.67%)
3.3	It was useful to see my daily pollution caused compared to my average pollution, as well as others average pollution (In percentage)	18	2 (11.11%)	2 (11.11%)	14 (77.78%)
3.4	It was useful to compare my energy consumption to the distance I could have been able to drive an electric car with the same energy consumption	18	3 (16.67%)	2 (11.11%)	13 (72.22%)
3.5	It was useful to compare my pollution to the distance I could have been able to drive a petrol car, to generate the same amount of pollution	18	3 (16.67%)	2 (11.11%)	13 (72.22%)
3.6	It was useful to get push notifications with pollution hints and updates	18	4 (22.22%)	5 (27.78%)	9 (50.0%)
3.7	It was useful to get reports on my pollution generated from Internet activity previous days	18	1 (5.56%)	4 (22.22%)	13 (72.22%)
3.8	It was useful to be able to explore a map to see where in the world I pollute	18	1 (5.56%)	3 (16.67%)	14 (77.78%)

3.9	It was useful to have a chart with the average usage of everyone else compared to my usage.	18	1 (5.56%)	3 (16.67%)	14 (77.78%)

Table 13.13: Results from part 3: Plugin Features

Rank	Feature	Score
1	It was useful to see my daily pollution caused compared to my average pollution, as well as others average pollution (In percentage)	75
1	It was useful to be able to explore a map to see where in the world I pollute	75
1	It was useful to have a chart with the average usage of everyone else compared to my usage.	75
2	It was useful to receive tips on how I can reduce my digital carbon footprint	73
2	It was useful to see my usage and pollution caused in real-time	73
2	It was useful to compare my pollution to the distance I could have been able to drive a petrol car, to generate the same amount of pollution	73
3	It was useful to get reports on my pollution generated from Internet activity previous days	72
4	It was useful to compare my energy consumption to the distance I could have been able to drive an electric car with the same energy consumption	71
5	It was useful to get push notifications with pollution hints and updates	62

Table 13.14: The features of the system, ordered by popularity.

Part 4: The Internet and the Environment

In Table 13.13, the results from Part 4 of Questionnaire 2 is shown. It shows that most of the participants know how the cloud works, and they also know how they can reduce their digital carbon footprint.

Part 5: The Environment

Table 13.16 displays the results from the 5th part of the second questionnaire. The results show that 4 participants feel guilty when using the Internet. It also indicates that most respondents do not think about the environment before visiting a website, posting to

ID	Question	n	D	N	A
4.1	I know how the "cloud" works	18	1 (5.56%)	4 (22.22%)	13 (72.22%)
4.2	I know how I can reduce my digital carbon footprint	18	2 (11.11%)	3 (16.67%)	14 (77.78%)

Table 13.15: Results from part 4: The Internet and the Environment

social media, or streaming videos online. However, 7 of the respondents think about the environment before uploading a file to the cloud.

ID	Question	n	D	N	A
5.1	I feel guilty about using the Internet.	18	6 (33.33%)	8 (44,344%)	4 (22.22%)
5.2	I think about the environment before I visit a website	18	10 (55.56%)	6 (33.33%)	2 (11.11%)
5.3	I think about the environment before I post an image to social media	18	9 (50.0%)	4 (27.78%)	4 (22.22%)
5.4	I think about the environment before I watch a video on the Internet	18	9 (50.0%)	7 (38.89%)	2 (11.11%)
5.5	I think about the environment before I stream a video	18	8 (44.44%)	6 (33.33%)	4 (22.22%)
5.6	I think about the environment before I upload a file to the cloud	18	7 (38.89%)	4 (22.22%)	7 (38.89%)

Table 13.16: Results from part 5: The Environment

Part 6: Final Questions

The result from *Question 6.1: Based on the feedback you got through the plugin, what do you experience with your own digital carbon footprint?*, is shown in Figure 13.12. It reveals that 76% of the participants experienced that their digital carbon footprint was greater than expected.

The results from question 6.2 to 6.4 is shown in Table 13.17. The results indicate that one-third think that the extension helped them reduce their digital carbon footprint, and 83% experienced that it helped them become more aware of how they pollute when using the Internet. Furthermore, 12 of the 18 respondents say that they better understand how the Internet works than before joining the project. However, seven say they have not changed their Internet habits after joining the project.

In Figure 13.13, the answers to *Question 6.2: Did you do any of the following to reduce your digital carbon footprint?* are presented. It shows that the participants did change their behavior on the Internet to reduce their digital carbon footprint. Four participants used the Internet less and turned down the quality when streaming video.

Six respondents deleted apps from their phones, while 12 respondents deleted files from the cloud.

The results from the final question *Number of files you have deleted from the cloud after joining the project* is shown in Figure 13.14 and indicates that 6 of the participants did not delete any files from the cloud. On the other side, 7 participants did delete more than 21 files from the cloud each.

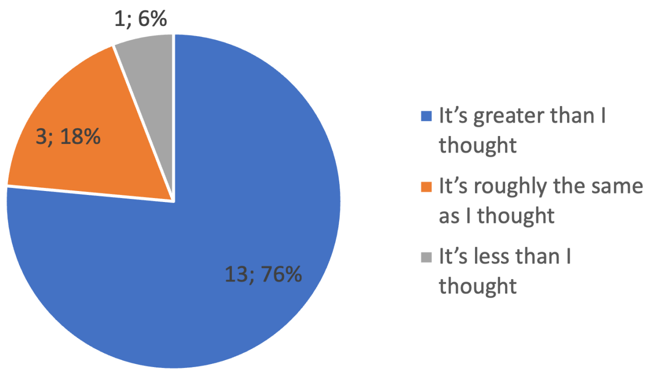

Figure 13.12: Results from Questionnaire 2, Question 6.1: Based on the feedback you got through the plugin, what do you experience with your own digital carbon footprint?

ID	Question	n	D	N	A
6.2	The plugin helped me reduce my digital carbon footprint	18	3 (16.67%)	9 (50%)	6 (33.33%)
6.3	The plugin helped me become more aware of how I pollute when I use the Internet	18	0 (0.0%)	3 (16.67%)	15 (83.33%)
6.4	I have a better understanding of how the Internet works	18	3 (16.67%)	3 (16.67%)	12 (66.67%)
6.5	I have changed my Internet habits	18	7 (38.89%)	6 (33.33%)	5 (27.78%)

Table 13.17: Results from part 6, questions 6.2-6.5: Final Questions

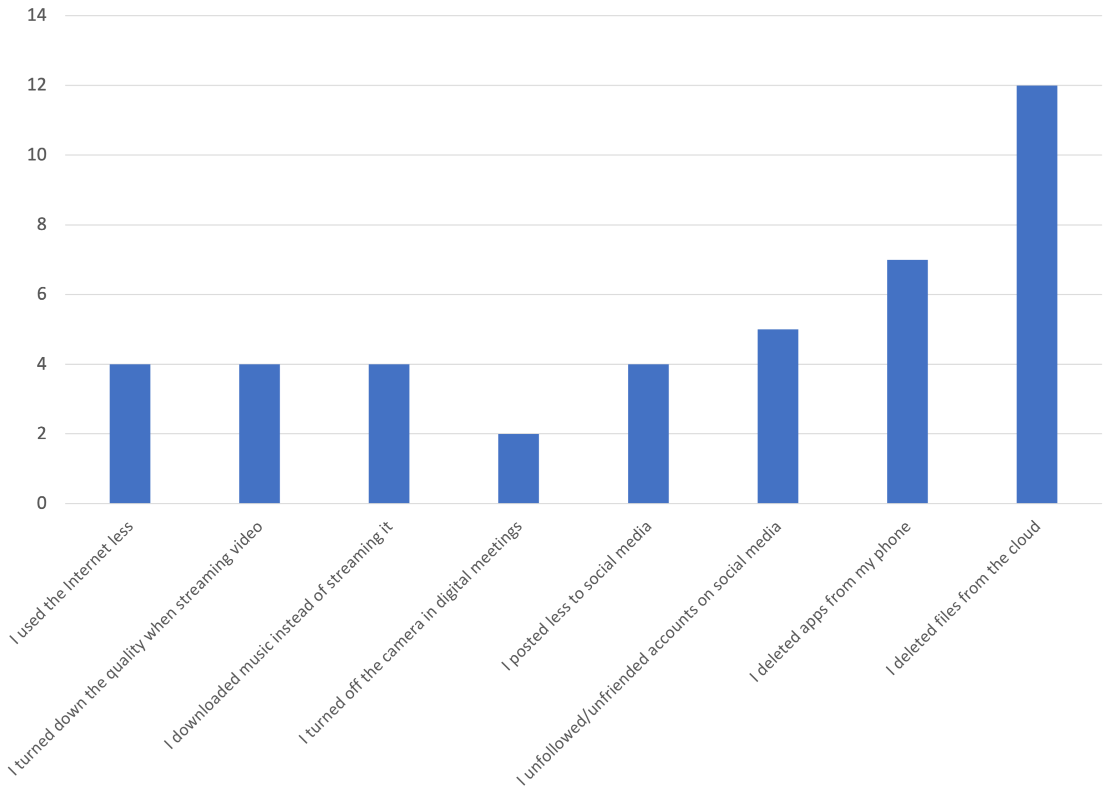

Figure 13.13: Results from Questionnaire 2, Question 6.2: Did you do any of the following to reduce your digital carbon footprint?

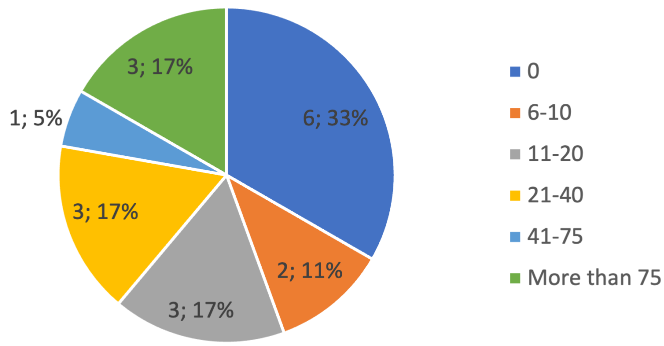

Figure 13.14: Results from Questionnaire 2, Question 6.3: Number of files you have deleted from the cloud after joining the project

13.2.3 Qualitative Feedback

The final question of the second questionnaire was an open question where the participants were asked if they wanted to give any feedback to the project. The result is shown in Table 13.18. Overall, the feedback on the project was positive. Feedback 8 and 9 indicate that their estimated carbon footprint was lower than their actual carbon footprint. Feedback 4, 6 and 10 shows that the participants got more aware of their digital carbon footprint after being part of the project.

ID	Feedback
1	Veldig bra plugin! Her har du vært på jobb, stå på med masterskrivinga :D
2	very interesting to be a part of this project. Very well done, Jesper!
3	This project is great! I really want to keep my plug-in and have it after the project is finished. I hope you will keep working with this and make a plug-in we can keep! :) Good luck!!!
4	This was super useful and gave me new insight and understanding of how much carbon and energy waste we create online, and what we can do ro reduce it
5	As a web developer it would be good if the plugin could be used to easy analyse the carbon footprint of a website across its many pages. I'm keen that the sites that I help produce are as energy efficient as possible.
6	Interesting idea to bring this awareness to the average user. I am sure most people don't think about their online activities at all in the context of pollution.
7	I thought the idea was very interesting, and it was great to see the metrics that had been tracked based on my usage. I found it hard to understand what they meant, and so the dashboard overall didn't feel as impactful as it might have been if there was an onboarding explanation of what the metrics mean. Since I didn't have a foundational understanding either, the contextual comparison didn't feel meaningful either. The measurement I understood the most was "Bytes", the others were much less meaningful if at all. It was insightful to see where in the world the services I was using were located in the world. It made me ponder the global reach of different online services.
8	My answers could be a bit misleading, as I was not able to have the plugin active on my work computer, where I spend 95% of my internett-surfing time (not including smart phone usage).
9	Brilliant approach and I'd love to continue to use this tool. Just to inform your data - my usage seems a great deal lower than most people, but I suspect it's because my browser use is spread between chrome and firefox so not an accurate picture of all my uses - but really inspired by your work and approach. Good luck with your dissertation and I hope to see a live version after that too. If I can help with that, please let me know.
10	Cool feedback, it's definitely a part of one's climate footprint which isn't talked a lot about. I really enjoyed using the plugin, and it was executed in a nice manner.

11	Really interesting project!

Table 13.18: Results from Questionnaire 2: Final feedback.

Age

When grouping the second questionnaire by age, only one result was borderline to significant. The question is displayed in Table 13.19. Of the 12 participants who answered this question (Participants who did not play online games prior to the project were told not to answer this question), nine were below 35, and 3 were 35 or older. All three above 35 disagreed with the statement, while 5 of the younger group were neutral.

ID	Question	Group	n	D	N	A	p
2.7	After joining the project, I spend less time on online gaming	< 35	9	4 (44.44%)	5 (55.56%)	0 (0.0%)	0.09407
		>= 35	3	3 (100.0%)	0 (0.0%)	0 (0.0%)	

Table 13.19: Results from the second questionnaire, grouped by age.

Gender

To make two groups and even out the sample size, others were added to the females. The results when dividing the participants by gender are shown in Table 13.20. The female group found the reports more interesting than the males. At the same time, 75% of the males did not think about the environment before entering a website, compared to only one female. Further, the females think more about the environment before posting an image to social media and streaming a video, compared to the males.

ID	Question	Group	n	D	N	A	p
3.7	It was useful to get reports on my pollution generated from Internet activity previous days	Male	12	1 (8.33%)	3 (25.0%)	8 (66.67%)	**0.03337**
		Female + Other	6	0 (0.0%)	1 (16.67%)	5 (83.33%)	
3.9	It was useful to have a chart with the average usage of everyone else compared to my usage.	Male	12	1 (8.33%)	2 (16.67%)	9 (75.0%)	0.06369
		Female + Other	6	0 (0.0%)	1 (16.67%)	5 (83.33%)	
5.1	I feel guilty about using the Internet.	Male	12	5 (41.67%)	6 (50.0%)	1 (8.33%)	0.06746
		Female + Other	6	1 (16.67%)	2 (33.33%)	3 (50.0%)	
5.2	I think about the environment before I visit a website.	Male	12	9 (75.0%)	2 (16.67%)	1 (8.33%)	**0.03388**
		Female + Other	6	1 (16.67%)	4 (66.67%)	1 (16.67%)	

5.3	I think about the environment before I post an image to social media.	Male	12	8 (66.67%)	3 (25.0%)	1 (8.33%)	**0.02283**
		Female + Other	6	1 (16.67%)	2 (33.33%)	3 (50.0%)	
5.4	I think about the environment before I watch a video on the Internet.	Male	12	8 (66.67%)	3 (25.0%)	1 (8.33%)	0.06819
		Female + Other	6	1 (16.67%)	4 (66.67%)	1 (16.67%)	
5.5	I think about the environment before I stream a video.	Male	12	7 (58.33%)	4 (33.33%)	1 (8.33%)	**0.03992**
		Female + Other	6	1 (16.67%)	2 (33.33%)	3 (50.0%)	

Table 13.20: Results from the second questionnaire, grouped by gender. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

Daily Time Spent Online

The results when dividing participants based on their answers on aggregated time spent online are shown in Table 13.21, based on the results presented in Section 13.2.1. The participants with an aggregated time lower than the median of 11.25 are in one group, and the rest is in the other group. Also, the participants in the lower group tend to spend less time on platforms like Facebook and more often turn off the camera in meetings than those who spend much time online. 89% of the participants with high usage know how the cloud works, while only 55% of the respondents with low usage answered agreeing to the same statement. Further, the participants with low Internet usage feel more guilty about using the Internet than the group that spends much time on the Internet.

ID	Question	Group	n	D	N	A	p
2.3	After joining the project, I spend less time on platforms like Facebook, LinkedIn, and Twitter	< 11.25	8	3 (37.5%)	2 (25.0%)	3 (37.5%)	0.08161
		>= 11.25	9	6 (66.67%)	2 (22.22%)	1 (11.11%)	
2.5	After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom.	< 11.25	9	3 (33.33%)	3 (33.33%)	3 (33.33%)	0.05263
		>= 11.25	8	6 (75.0%)	2 (25.0%)	0 (0.0%)	
2.6	After joining the project, I spend less time on regular browsing	< 11.25	9	1 (11.11%)	5 (55.56%)	3 (33.33%)	0.07155
		>= 11.25	9	4 (44.44%)	4 (44.44%)	1 (11.11%)	
4.2	I know how the "cloud" works.	< 11.25	9	1 (11.11%)	3 (33.33%)	5 (55.56%)	0.06009
		>= 11.25	9	0 (0.0%)	1 (11.11%)	8 (88.89%)	
5.1	I feel guilty about using the Internet.	< 11.25	9	1 (11.11%)	5 (55.56%)	3 (33.33%)	**0.04511**
		>= 11.25	9	5 (55.56%)	3 (33.33%)	1 (11.11%)	
5.2	I think about the environment before I visit a website.	< 11.25	9	4 (44.44%)	3 (33.33%)	2 (22.22%)	0.09402
		>= 11.25	9	6 (66.67%)	3 (33.33%)	0 (0.0%)	
5.4	I think about the environment before I watch a video on the Internet.	< 11.25	9	3 (33.33%)	4 (44.44%)	2 (22.22%)	0.08553
		>= 11.25	9	6 (66.67%)	3 (33.33%)	0 (0.0%)	

Table 13.21: Results from the second questionnaire, grouped by time spent on the Internet. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

Experience with Developing/Designing Websites

Table 13.22 shows the results when dividing the group based on their experience with developing or designing websites or apps that use the Internet. While the experienced participants tend not to think about the environment before visiting a website, the non-experienced ones do. Similarly, a larger share of the non-experienced think about the environment before posting an image to social media, watching a video on the Internet, or streaming a video (On streaming platforms like Netflix), compared to the non-experienced ones. Further, the participant with no experience found the push notification feature of the final application more valuable than the participants with experience.

ID	Question	Group	n	D	N	A	p
3.6	It was useful to get push notifications with pollution hints and updates	None	7	1 (14.29%)	1 (14.29%)	5 (71.43%)	**0.04054**
		Experi.	11	3 (27.27%)	4 (36.36%)	4 (36.36%)	
4.3	I know how I can reduce my digital carbon footprint.	None	7	0 (0.0%)	1 (14.29%)	6 (85.71%)	0.08791
		Experi.	11	2 (18.18%)	2 (18.18%)	7 (63.64%)	
5.2	I think about the environment before I visit a website.	None	7	0 (0.0%)	5 (71.43%)	2 (28.57%)	**0.00059**
		Experi.	11	10 (90.91%)	1 (9.09%)	0 (0.0%)	
5.3	I think about the environment before I post an image to social media.	None	7	1 (14.29%)	2 (28.57%)	4 (57.14%)	**0.00496**
		Experi.	11	8 (72.73%)	3 (27.27%)	0 (0.0%)	
5.4	I think about the environment before I watch a video on the Internet.	None	7	1 (14.29%)	5 (71.43%)	1 (14.29%)	**0.03516**
		Experi.	11	8 (72.73%)	2 (18.18%)	1 (9.09%)	
5.5	I think about the environment before I stream a video.	None	7	0 (0.0%)	5 (71.43%)	2 (28.57%)	**0.02317**
		Experi.	11	8 (72.73%)	1 (9.09%)	2 (18.18%)	

Table 13.22: Results from the first questionnaire, grouped by their experience with developing and/or designing websites and/or apps that use the Internet. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

Willingness to Change

In Table 13.23, the results from dividing the participants by their answer on question 6.1 in the first questionnaire are presented. The participants that answered neutral or disagreeing to the statement *I'm willing to change my behavior on the Internet to save the environment* are grouped in the "Low willingness" group, while the participants that answered agreeing to the statement is placed in the "High willingness" group. The results show that the participants with high willingness found the generated reports and charts in the application more valuable than those with low willingness. It also reveals that the respondents with low willingness think less about the environment when visiting a website, posting to social media, and watching a video, compared to the users with high willingness.

ID	Question	Group	n	D	N	A	p
2.4	After joining the project, I spend less time on streaming platforms like Netflix and HBO Max	Low	7	5 (71.43%)	1 (14.29%)	1 (14.29%)	0.06032
		High	9	3 (33.33%)	3 (33.33%)	3 (33.33%)	
2.5	After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom	Low	7	5 (71.43%)	2 (28.57%)	0 (0.0%)	0.07776
		High	10	4 (40.0%)	3 (30.0%)	3 (30.0%)	
3.7	It was useful to get reports on my pollution generated from Internet activity previous days	Low	8	1 (12.5%)	2 (25.0%)	5 (62.5%)	**0.01667**
		High	10	0 (0.0%)	2 (20.0%)	8 (80.0%)	
3.8	It was useful to be able to explore a map to see where in the world I pollute	Low	8	1 (12.5%)	2 (25.0%)	5 (62.5%)	0.05412
		High	10	0 (0.0%)	1 (10.0%)	9 (90.0%)	
3.9	It was useful to have a chart with the average usage of everyone else compared to my usage.	Low	8	1 (12.5%)	2 (25.0%)	5 (62.5%)	**0.01982**
		High	10	0 (0.0%)	1 (10.0%)	9 (90.0%)	
5.3	I think about the environment before I visit a website.	Low	8	7 (87.5%)	0 (0.0%)	1 (12.5%)	**0.02182**
		High	10	3 (30.0%)	6 (60.0%)	1 (10.0%)	
5.3	I think about the environment before I post an image to social media.	Low	8	6 (75.0%)	2 (25.0%)	0 (0.0%)	**0.01489**
		High	10	3 (30.0%)	3 (30.0%)	4 (40.0%)	

5.4	I think about the environment before I watch a video on the Internet.	Low	8	7 (87.5%)	1 (12.5%)	0 (0.0%)	**0.00531**
		High	10	2 (20.0%)	6 (60.0%)	2 (20.0%)	
5.5	I think about the environment before I stream a video.	Low	8	6 (75.0%)	2 (25.0%)	0 (0.0%)	**0.00935**
		High	10	2 (20.0%)	4 (40.0%)	4 (40.0%)	

Table 13.23: Results from the second questionnaire, grouped by their willingness to change their behaviour to save the environment. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

Estimated CO_2e Emissions

In Table 13.24, the results from dividing the participants by their estimated CO_2e emissions are presented. The numbers are fetched from the database. Participants with a total estimated CO_2e emission lower than the median of 168.84 grams are placed in one group, and the rest is placed in the other group. The results show that participants with low usage more often turn off their cameras in digital meetings. They also found the tips more valuable than the group with high usage. 8 of the 9 participants with low usage answered agreeing to the statement "*I know how I can reduce my digital carbon footprint.*", while only 55% of the respondents with high usage did the same. In comparison, the participants with high usage feel less guilty when using the Internet than those with low usage.

ID	Question	Group	n	D	N	A	p
2.3	After joining the project, I spend less time on platforms like Facebook, LinkedIn, and Twitter	Low	8	3 (37.5%)	2 (25.0%)	3 (37.5%)	0.08161
		High	9	6 (66.67%)	2 (22.22%)	1 (11.11%)	
2.5	After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom	Low	9	3 (33.33%)	3 (33.33%)	3 (33.33%)	**0.01707**
		High	8	6 (75.0%)	2 (25.0%)	0 (0.0%)	
3.1	It was useful to receive tips on how I can reduce my digital carbon footprint	Low	9	0 (0.0%)	1 (11.11%)	8 (88.89%)	**0.01968**
		High	9	1 (11.11%)	3 (33.33%)	5 (55.56%)	

3.2	It was useful to see my usage and pollution caused in real-time	Low	9	0 (0.0%)	1 (11.11%)	8 (88.89%)	**0.01733**
		High	9	2 (22.22%)	3 (33.33%)	4 (44.44%)	
3.5	It was useful to compare my pollution to the distance I could have been able to drive a petrol car	Low	9	0 (0.0%)	1 (11.11%)	8 (88.89%)	0.08667
		High	9	3 (33.33%)	1 (11.11%)	5 (55.56%)	
4.3	I know how I can reduce my digital carbon footprint.	Low	9	0 (0.0%)	1 (11.11%)	8 (88.89%)	**0.04925**
		High	9	2 (22.22%)	2 (22.22%)	5 (55.56%)	
5.2	I feel guilty about using the Internet.	Low	9	1 (11.11%)	5 (55.56%)	3 (33.33%)	**0.04511**
		High	9	5 (55.56%)	3 (33.33%)	1 (11.11%)	
5.3	I think about the environment before I watch a video on the Internet.	Low	9	2 (22.22%)	5 (55.56%)	2 (22.22%)	**0.01486**
		High	9	7 (77.78%)	2 (22.22%)	0 (0.0%)	

Table 13.24: Results from the second questionnaire, grouped by their estimated CO_2 emissions, fetched from the database. Showing questions where the probability value (p) is significant ($p < 0.5$, highlighted in bold) and borderline to significant values ($p < 0.1$).

13.2.4 Results From the Questionnaires Compared

To compare the results from both questionnaires, the Wilcoxon Signed Rank Test is used. The results from comparing the results are shown in Table 13.25, and the significant results are highlighted in bold. Only participants that answered both questionnaires are included in the results.

The results show that the participants feel more guilty when using the Internet after attending the project. In comparison, it can not be concluded that the participants know more about the cloud than before the project. However, the rest of the results are significant. Thus, the participants think more about the environment when using different Internet services than before the project.

ID	Question	Group	n	D	N	A	p
5.1 - 5.2	I feel guilty about using the Internet.	Before	18	15 (83.33%)	3 (16.67%)	0 (0.0%)	**0**
		After	18	6 (33.33%)	9 (50.0%)	3 (16.67%)	
4.2 - 4.2	I know how the "cloud" works	Before	18	4 (22.22%)	1 (5.55%)	13 (72.22%)	15
		After	18	1 (5.56%)	4 (22.22%)	13 (72.22%)	
4.5 - 4.3	I know how I can reduce my digital carbon footprint	Before	18	13 (72.22%)	3 (16.67%)	2 (11.11%)	**0**
		After	18	2 (11.11%)	3 (16.67%)	13 (72.22%)	
5.3 - 5.2	I think about the environment before I visit a website	Before	18	18 (100.0%)	0 (0.0%)	0 (0.0%)	**0**
		After	18	10 (55.56%)	6 (33.33%)	2 (11.11%)	
5.4 - 5.3	I think about the environment before I post an image to social media	Before	18	18 (100.0%)	0 (0.0%)	0 (0.0%)	**0**
		After	18	10 (55.56%)	4 (22.22%)	4 (22.22%)	
5.5 - 5.4	I think about the environment before I watch a video on the Internet	Before	18	18 (100.0%)	0 (0.0%)	0 (0.0%)	**0**
		After	18	10 (55.56%)	6 (33.33%)	2 (11.11%)	
5.6 - 5.5	I think about the environment before I stream a video.	Before	18	0 (0.0%)	1 (11.11%)	8 (88.89%)	**0**
		After	18	8 (44.44%)	6 (33.33%)	4 (22.22%)	
5.7 - 5.6	I think about the environment before I upload a file to the cloud.	Before	18	17 (94.44%)	0 (0.0%)	1 (5.56%)	**0**
		After	18	6 (33.33%)	5 (27.78%)	7 (38.89%)	

Table 13.25: Results from the first and second questionnaire, compared. Questions where the probability value (p) is significant is highlighted in bold.

13.3 Results From the DataCollector

This section will present the results from the data collected through the chrome extension. In total, 20 participants installed the extension. The data from these two weeks will be illustrated using line and box diagrams. During the first week of the experiment, the participants did not get feedback on their usage. After the first week, the participants were able to explore their results. The red horizontal dotted line in the diagrams indicates the day the users got access to see their usage. Four participants are excluded from the results, as they installed the extension after the seventh day. The data was collected for the following attributes:

- **Size**: Size in grams
- **CO_2**: Estimated CO_2e emissions (Based on destination country for each network request)
- **kWh**: Estimated kWh (Proportional to the size)
- **Number of calls**: Number of network requests
- **Seconds active**: Seconds the user was active

All the diagrams for each attribute are presented in Appendix 8.

13.3.1 Usage Per User

This section will show the usage per user for the following attributes: Size, Estimated CO_2 emissions, and Seconds active. Line diagrams will be presented for each of the attributes. The line diagram is chosen to better understand the different participants' Internet usage throughout the project. Each color represents a participant, and the same color is used for the same participant throughout all the line diagrams.

Size

The line diagram in Figure 13.15 shows the total size of bytes sent and received per user per day. The diagram shows a significant variation in the amount of data used per participant, especially after day 7. In total, no user had exceptionally high usage every day.

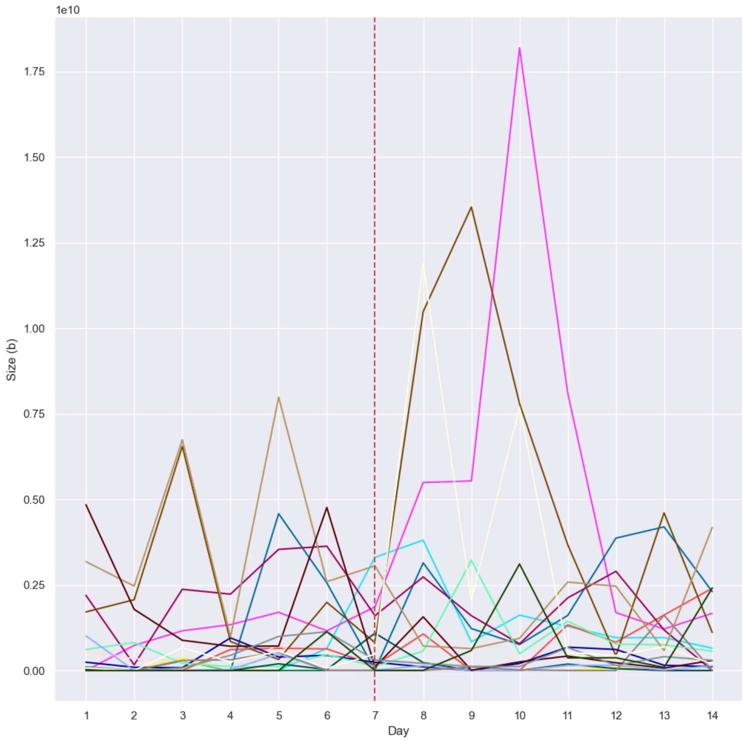

Figure 13.15: Line diagram for each user, showing how the size varied per user. The red dotted line indicates when the users got access to see their usage.

Estimated CO_2e Emissions

Figure 13.16 presents the estimated CO_2 emissions per user per day. Before day 7, the amount of CO_2 emissions was between 0 and 160 g. However, on day eight, the emissions from most of the users increased. After that, three participants (Pink, Brown, and Turquoise) increased their usage further. The emissions from the rest of the participants decreased until days 9 and 10, where they increased slightly again.

Figure 13.16: Line diagram for each user, showing how the estimated CO_2e varied per user. The red dotted line indicates when the users got access to see their usage.

Seconds Active

The line diagram in Figure 13.17 shows the seconds active per user. This attribute is relevant to the thesis, as the seconds active indicate the energy used by the end-user device. Indications of a trend can be seen; the seconds active increased until the eighth day before it decreased for some days and then started to increase again. One user was significantly more active than the rest of the participants.

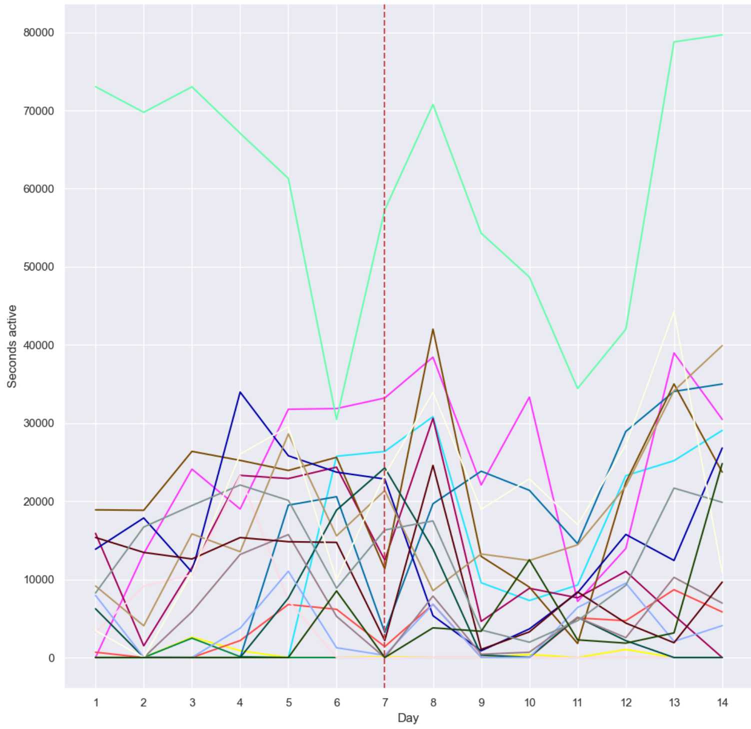

Figure 13.17: Line diagram for each user, showing seconds active per day per user. The red dotted line indicates when the users got access to see their usage.

13.3.2 Box Diagrams

In this section, box diagrams for the attributes *Size*, CO_2, and *Seconds Active* will be shown. The box diagram is used to give a visual indication of the distribution of data per day, making it more manageable to look for trends. The chosen diagrams include outliers, with the boundaries of the whiskers at 1.5 times the distance of the Interquartile Range. As a single participant can considerably impact the median of the results, outliers were chosen when drawing the plots to minimize the effect these had on the trends.

Size

In Figure 13.18, the box diagram for the amount of data used per day per user is presented. Until day seven, the amount of data varies. However, a substantial increase in data usage is seen from day seven to day eight. This is the most significant increase in data between two days during the project. The median of the box at day 8 is similar to the median seen at day 6, but the third quartile is more prominent than for any other day. Furthermore, days 8-11 and 14 have a higher third quartile and maximum than 6 of the days of the first week. Similarly, more outliers are observed in the second week than in the first week. These observations indicate a higher usage the second week than the first week.

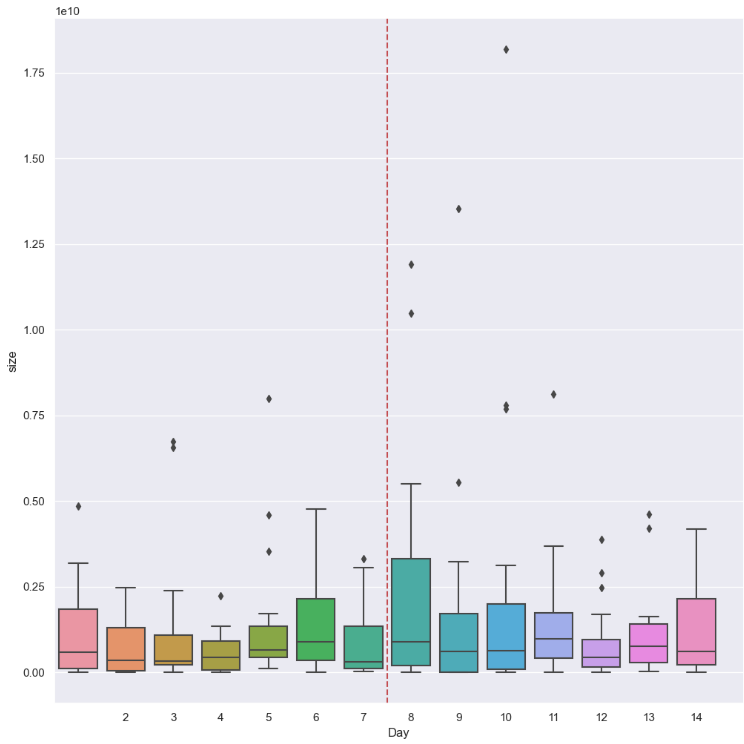

Figure 13.18: Box diagram, showing the amount of data sent each day. The red dotted line indicates when the users got access to see their usage.

Estimated CO_2e Emissions

Figure 13.19, shows the box diagram for the estimated CO_2e emissions per user per day. A remarkable difference in estimated emissions is seen from day 7 to day 8. This diagram also presents several outliers, with the highest recorded value being roughly 20 times as large as the median for the same day. Also, the average size of the outliers in the second week is greater than the average size of the outliers in the first week. However, it is not a significant trend in the data, but there is an indication of higher estimated emissions from day 8-14 compared to day 1-7.

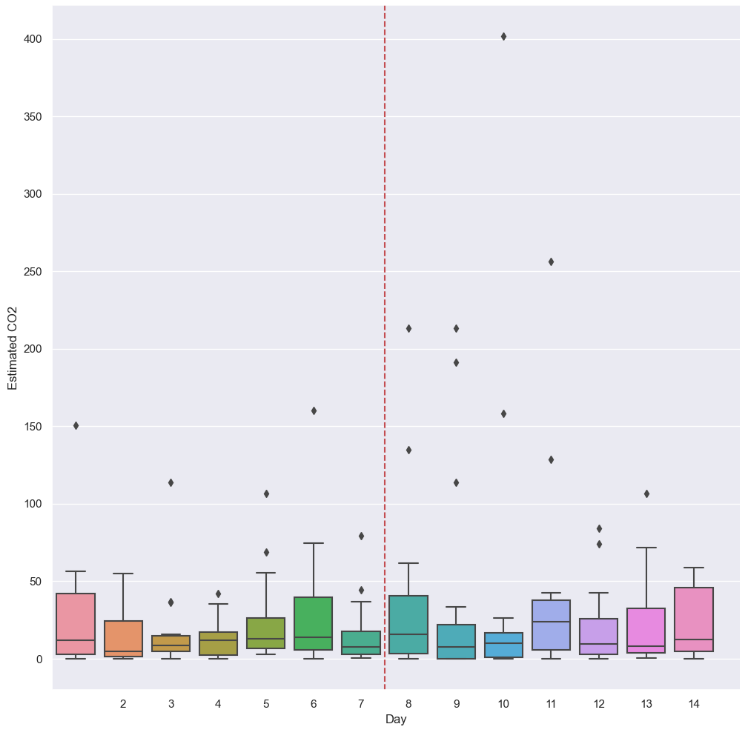

Figure 13.19: Box diagram, showing the amount of the estimated CO_2e emissions per day. The red dotted line indicates when the users got access to see their usage.

Seconds Active

The box diagram in Figure 13.20 shows the seconds active per user per day. When looking at the third quartiles, there is an increase in seconds active from day 1 to day 5, before a slight decrease until day 7. From day 7 to day 8, there is an increase in both the median, the 75th quartile, and the maximum. However, the lowest median during the experiment is observed at day eight. From day 9 to 11, the median indicates that the participants used their devices less than any other period during the project.

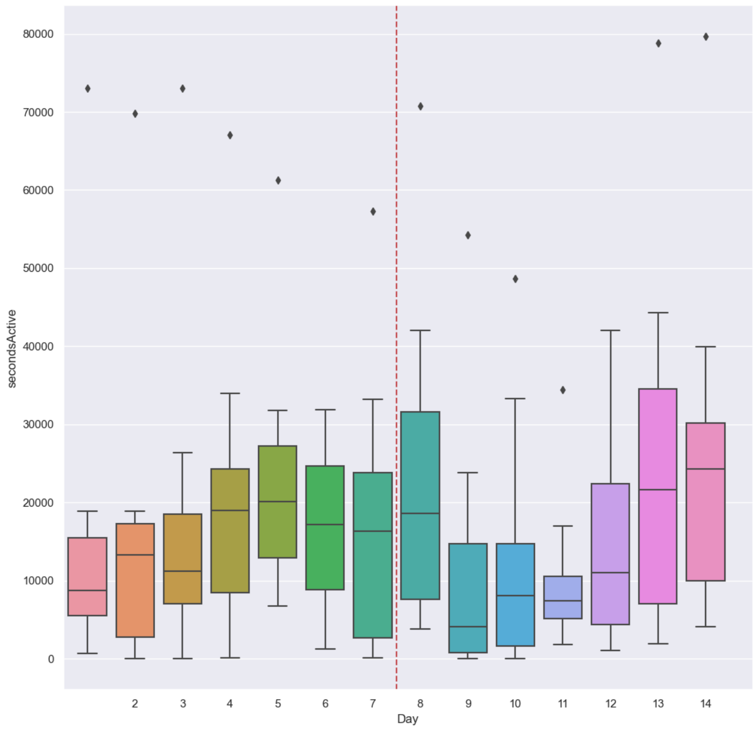

Figure 13.20: Box diagram, showing seconds active per day. The red dotted line indicates when the users got access to see their usage.

13.3.3 Means of Means

The means of means are presented in Table 13.26. The means of means is presented to explore the difference in usage between the two weeks of the experiment. The data collected from the first week is in one group, and the data from the second week is in the other group. The difference between the means is shown in the final column. A negative number indicates that the usage was higher the second week than the first week.

Out of the five attributes, four indicate a higher usage the second week. However, the means of the seconds the users were active the second week were lower than the first week. This indicates that the number of bytes sent during the second week happened over a shorter period compared to the first week.

Attribute	**First week**	**Second week**	**First week - Second week**
Size	1111194528	1713449598	-602255070
Estimated CO2	20.93	32.53	-11.59
kWh	0.09313	0.14361	-0.05048
Number of calls	22148	22148	-22148
Seconds active	17270	16717	**553**

Table 13.26: Mean of means for each of the attributes, divided into two groups.

13.4 Summary

In this chapter, the results from the data generation methods have been presented.

First, the results from the two questionnaires were presented. Then the results were grouped based on Age, Gender, Hours used daily on the Internet, Willingness to change, and experience with developing or designing websites or apps. These groups were then used to look for significant differences. The results show that people with low willingness to change or high Internet usage think less about the environment when performing tasks on the Internet.

Second, the results from the identical questions of the two questionnaires were also analyzed to look for significant values. The results showed that the participants, to a larger extent, thought about the environment before performing various tasks on the Internet, compared to before they joined the project. Most of the participants also found the project helpful to become more aware of their digital carbon footprint.

Finally, the data collected from the chrome extension have been shown using line and box diagrams. No clear trends could be seen in the diagrams, but there are indications that the amount of data used increased after the participants got access to see their usage. This finding is strengthened when analyzing the difference of the means of means.

14 Discussion

This chapter aims to discuss the findings and results of this thesis in light of the research questions. Further, the limitations of this thesis will be given. In the end, the proposals for future work will be presented.

14.1 Discussing the Research Questions

14.1.1 RQ 1: What is the carbon footprint of the Internet today and what are the trajectories for the future?

An essential part of this thesis was to understand the Internet's carbon footprint, based on a literature review of the research conducted in this field. As presented in Chapter 5, there is still a significant uncertainty when it comes to the Internet and its carbon footprint.

For this project, a factor of 0.09 kWh/GB was used, based on the research conducted by Andrae [6] and Obringer et al. [110]. The research executed by Obringer et al. did not provide an actual estimation for kWh/GB, but a median carbon footprint of 32.13 g CO_2/GB. The world carbon footprint was then used to convert their estimations to 0.068 kWh/GB. On one side, this number is much higher than the numbers reported by Malmodin and the IEA. On the other side, this number is lower than both the estimates of the ShiftProject and Ruiz et al.

There is not a single answer to what the Internet's carbon footprint is today. However, the rest of this section will elaborate on why there are variations, why the factor of 0.09 GB/kWh was chosen, and what the future of the Internet's carbon footprint looks like.

Public Data Billstein et al. eight identified challenges are of good help when discussing the significant variation in estimations. From the literature review conducted, I argue that the main problem for the ICT sector is the limited available data. Life Cycle Analysis (LCA) used when estimating the Internet's carbon footprint is not specific to the Internet; these calculations are used for products across all sectors. However, the unique thing for the ICT sector is the limited publicly available data. When conducting an LCA of a piece of clothing, the materials and energy required are easier to track, as the concept of the clothing is a limited amount of measurable and visual materials. When it comes to the LCA of the Internet, the same visual measurable units do not exist. Different stakeholders have information about different parts of the Internet. While getting information about the energy used in the production of Internet components might be possible, the amount of data that travels around the world and the energy that it uses are harder to dig up. The only way to access these data is through the Internet Service Providers (ISP) who do not have any obligations to give these data to the public. The less available data, the larger the chance for estimations that do not replicate the real world.

On the other hand, it can be practically impossible to make the amount of data that travels through the Internet and the energy it uses available to the public. Therefore, thorough research needs to be conducted on how to make a system that the ISPs can use to capture and report the amount of data sent across the Internet and its resulting energy usage and carbon footprint. The data collected by the system should then be made publicly available, ensuring that more research can be accomplished in the field with less uncertainty about the data.

Today, researchers that have access to these data work in companies that have an incentive not to damage the ICT sector's reputation, such as Ericsson (Malmodin) and Huawei (Andrae). Thus, having a neutral organ responsible for following up the numbers reported by the ISPs can benefit research in this field.

Water- and Land Footprint As mentioned by Billstein et al., the water- and land footprint of the ICT sector is a challenge to research further. Obringer et al. did in their research find that the world water footprint for transmission and storage varies from 0.15 L/GB to 34.99 L/GB, with a median of 0.74 L/GB. They also found that the world land footprint was from 0.68 cm^2/GB to 19.98 cm^2/GB, with a median of 11.07 cm^2/GB [110]. They estimate that the Internet emits 97 million tons of CO^2e, has a water footprint of 2.6 trillion liters, and a land footprint of 3400 km^2 per year. These numbers indicate that including the water and land footprint in the environmental footprint of the Internet is essential. Due to the limited available data, it is complicated to verify these numbers. However, if the numbers are correct, the ICT sector needs to reduce more than just its carbon footprint.

The End of Moore's Law Another factor that can impact the future development of the Internet's carbon footprint is the end of Moore's Law. Moore's law observes that the price per performance factor rate will be halved every two years, getting more computational power for the same cost. Research shows that the observation described by Moore's law is slowing down [54]. The energy potential of digital components is getting smaller as technology is reaching its physical limits. This means that both the cost and energy usage might increase in the future [54]. It also indicates that the embodied emissions can increase as more energy is used when creating the Internet components.

On one side, the increased cost can slow down the production of new network components, restricting the carbon footprint of the ICT sector. Nevertheless, it might not slow down the development of the Internet but instead increase the cost for the end-user. If the costs for the ISPs and data centers increase, the end-user will have to cover the costs one way or the other. However, this can slow down the Internet's carbon footprint. Suppose Internet users are offered subscriptions based on either Internet Speed (Mb/s) or data traffic per month. In that case, end-users might prefer to choose lower-tier subscriptions, therefore being forced to limit the amount of data they are using, which will result in less data being transmitted across the Internet for a given time frame.

More People and Data It is expected to be more people connected to the Internet and more data sent across the Internet in the future.

As mentioned by Ruiz et al., the 2030 UN Agenda for Sustainable Development aims at providing access to the Internet in all developing countries. This results in the need for more Internet components and more data being transmitted across the world. Thus, the carbon footprint of the Internet will increase.

Both individuals and organizations are producing more and more data each year [30, 84]. Even though the efficiency of the components is going up, new components need to be added to the Internet to process the increased amount of data, which will increase the carbon footprint of the ICT sector.

Data Centers As discovered by Billstein et al., data centers are often excluded by the Life Cycle Analysis of the Internet. This can significantly impact the estimations of the Internet's carbon footprint. Koot et al. find in their simulations that the energy consumption by data centers in 2030 to be 321 TWh [81]. When they add the uncertainty from the future of Moore's law to their simulations, they estimate data centers to use 658 TWh.

However, seeing two of the major data center providers aiming at carbon neutrality is believed to reduce the carbon footprint of the ICT sector. Google and Microsoft are focusing on becoming carbon-free, creating a sustainable cloud future. Microsoft has said that 100% renewable energy should run Azure (Their cloud services) by 2025, and they should be water positive by 2030 [93]. Google, on their side, aims at operating on carbon-free energy 24/7 everywhere by 2030 and replenish 120% of the water they use by 2030 [64]. Their goal of replenishing more water than they use is also believed to reduce the water footprint of the ICT sector.

The Larger ICT Companies Benefit From Individuals Consuming and Creating More Data As mentioned by Gerry McGovern in the interview presented in Chapter 7, both social media platforms and cloud providers earn more money the more data the user creates and consumes. If the data collected by the social media platforms and the data stored in the cloud by individuals continue to rise, the Internet's carbon footprint is expected to rise. Thus, governmental legislation seems necessary to reduce the carbon footprint of large ICT companies. As long as the companies have no initiative to change, they will most likely not change, as noted by McGovern.

Uncertainty Aside, Reduction in Emissions Seems to be Necessary Even though there is substantial uncertainty around the carbon footprint of the Internet, research indicates that measurements must be made to make sure the ICT sector is on track with the IPPCs 1.5 °C trajectory [139, 87, 56]. The International Telecommunication Union concluded that the ICT sector needs to be reducing its emission by 51% by 2030 compared to 2015. Even though companies like Microsoft and Google aim to become carbon-free by 2025 and 2030, we should all have the responsibility to produce and consume less, both data and digital devices, if we want to contribute.

14.1.2 RQ 2: How can an application used to estimate the digital carbon footprint of a user be created without significantly increasing the digital carbon footprint of the user?

Using a web browser extension to estimate the digital carbon footprint introduced some challenges. As mentioned in Chapter 6, the Google Chrome API did not directly provide all the information needed to calculate the size and destination of the network requests, and some workarounds had to be implemented. However, the system created does show that it is possible to create a solution that estimates the digital carbon footprint of the Internet if simplifications are made to ensure that the formulas used to estimate the carbon footprint depends on the data collected, only.

On the other hand, it is believed that a desktop application is better suited to capture the digital carbon footprint of a user than a browser extension. When using a desktop application, it will be possible to capture the network requests from all applications running on the computer and the energy usage from the computer. Further, this data could be combined, resulting in more precise estimations of a user's carbon footprint, given that mathematical formulas to use these numbers are created.

This application used an API and a database to log the usage of the users and support gamification. Adding these components increased both the solution's embodied and operational carbon footprint. However, based on the answers from Questionnaire 2, Part 3: Plugin features, two of the highest scored features had gamification functionality. This indicates that the application's success is dependent on the gamification features of the application. In that case, the API and databases components should not be shelved.

Regardless of the implementation methods chosen for future applications, it is believed that the code and concepts developed during this project will benefit and aid researchers in the future.

The following goals when developing applications to estimate the carbon footprint of Internet users (and applications in general) are proposed:

- **Minimize the number of components in the system**
- **Minimize the size of the component(s)**
- **Minimize the data stored**
- **Minimize the calculations**
- **Minimize the data transmitted**
- **Minimize the embodied emissions from the developing phase**

Number of Components If the system is created using multiple components running in different environments, the total carbon footprint of the system will be larger than the carbon footprint of a single component system. This system used a cloud-hosted database and API in order to be able to monitor the usage of the participants stored. If the system will not need to monitor the usage of its users or persistent storage is not needed, both the database and API can be excluded from the final system. This means that the final system can consist of only one application running on the end-user device. As long as the final application does not use a lot of the end-user device's resources, it will lower its carbon footprint. Suppose the final application uses a lot of the end-user device's resources and the carbon intensity of the energy used to power the end-user device is high. In this case, it will most likely be better to move the calculations to a cloud service running on green power.

Size The size of the system will impact the carbon footprint of the solution. If the total size of the system is large, more data will be used when downloading, uploading, or updating the component(s) of the system. This will increase the embodied and operational carbon footprint of the system. Therefore, developers should aim at optimizing the size of the final system. Minimizing the use of videos, images, and animations reduces the carbon footprint of the end-user. The first questionnaire results show that 48% of the participants feel images, videos, and animations are essential to make web pages more attractive. However, 44% are willing to stop visiting websites that look less "fancy" and use fewer images and videos if it will reduce their digital carbon footprint (Question 6.7 and 6.8). Thus, minimizing the use of videos, images, and animations is not believed to reduce the number of users of an application and will instead aid in reducing the carbon footprint of the end-users using the system. The system developed for this thesis used modern JavaScript bundlers to reduce the size of the system.

Storage, Calculations and Data Transmitted The storage and calculation strategy is tightly coupled and will impact the amount of data that is transmitted. These attributes will, in turn, impact the carbon footprint of the users of the system. For this system, the data was aggregated directly daily and totally. This ensured that the data presented to the user only had to be aggregated once, thereby reducing the number of calculations and data stored. If the data had been stored directly, it would have needed to be aggregated every time the user wanted to explore its data. This would lead to significantly more calculations, more database reads and writes, and more data being stored, which again would increase the system's carbon footprint. However, it is challenging to correct wrong calculations when storing aggregated data. Therefore, high test coverage of the calculation methods is needed to reduce bugs.

Developers should optimize the calculation methods as much as possible. A calculation that takes a long time to complete or uses much computing power when executed will use more energy than a calculation that takes less time or requires less computational power.

Embodied Emissions From the Developing Phase Developers should aim at reducing the embodied emissions from the development of the system. Time spent developing the system will increase the embodied emissions of the system, as more energy will be used to power the electronic devices used to develop the system. Further, building projects often require a significant amount of computational power. Therefore, reducing the number of times the project is built will positively impact the system's carbon footprint. Developers should also try to minimize the amount of data that is downloaded during the development of the project, and using centralized package management strategies can reduce the amount of data being downloaded when developing the system. This project used pnpm to manage the third-party packages, which reduced the number of times packages had to be downloaded from the Internet.

14.1.3 RQ 3: How can people become aware of their digital carbon footprint?

By developing a system that visualizes a user's Internet usage, I have verified how a system can be developed to make users more aware of their digital carbon footprint.

41 of the 43 respondents that answered the first questionnaire answered yes on *Question 3.1: Before entering this study - did you know that the Internet pollutes?* Based on the number of respondents that answered yes to this question, it can be derived that people understand that they pollute when using the Internet. However, it is not given that they are aware of their own digital carbon footprint. This finding is supported by looking at the results from Questionnaire 1: *Question 4.4: I know how I can reduce my digital carbon footprint.* and *Question 4.5: I want to learn how I can reduce my digital carbon footprint.* Only 16% answered agreeing to Question 4.4, and 95% answered agreeing to Question 4.5. The results from these questions show that people are not aware of how they can reduce their digital carbon footprint. Hence, they do not know how their actions on the Internet affect their digital carbon footprint.

After the project, 83 % answered that the plugin helped them become more aware of how they pollute when using the Internet. This shows the effect the system had on the participants becoming more aware of their digital carbon footprint. Further, the results from the Question *I know how I can reduce my digital carbon footprint* demonstrate how the participants became aware of how they can reduce their digital carbon footprint. Of the 18 respondents who answered both questionnaires, 13 disagreed with this statement before the project. After the project, 13 agreed with the statement. Thus, people using the system became more aware of reducing their digital carbon footprint. This implies that people got more aware of their digital carbon footprint when using the system.

As discussed, there is uncertainty about the actual carbon footprint of Internet usage. 76% of the respondents answered that they experienced that their digital carbon footprint was greater than they thought. However, the results from the question mentioned can be influenced by the estimations presented to the participants. Suppose the numbers used to estimate the participant's carbon footprint is higher than the actual CO_2/GB of the Internet. In that case, it can have made more of the respondents answer that they found their digital carbon footprint to be greater than expected. This can have impacted the results of the second questionnaire.

14.1.4 RQ 4: Will people that become more aware of their digital carbon footprint change their behavior on the Internet?

This thesis's findings indicate that people who get more aware of their digital carbon footprint do not necessarily change their behavior on the Internet.

When analyzing the data that was collected through the plugin, it is not easy to see any clear trends in the data. However, the day after the participants got access to see their estimated emissions in real-time, they sent more data across the Internet than any other day. On one side, this indicates that users that get access to see their estimated carbon footprint will increase the data sent across the Internet. Nonetheless, the increased data usage can also be due to the users wanting to understand where in the world they pollute when using such services. Therefore they increase the amount of data they send across the Internet.

In total, more data was sent the week after the participants got access to see their data, compared to the week before. Nevertheless, they spent less time on their computer the second week than the first week. It is believed that it is necessary to log the usage for a longer time, with more participants, in order to draw any conclusions from the collected data. The results from the final feedback also indicate that some of the users did use other browsers to access the Internet, resulting in variations in their estimated emissions.

The questionnaires' results show that the participants think more about the environment before performing specific actions on the Internet compared to before they joined the project. 72% of the participants know how to reduce their digital carbon footprint after being part of the project. However, the second questionnaire results also show that most participants still spend as much time using different services before joining the project, indicating that they did not change their behavior. On the other hand, 7 of the participants answered agreeing to one or more of questions 2.1 to 2.7 (Information about your Internet usage). Thus, it can be reasoned that some of the participants changed their behavior during the project.

Based on the analysis from the data collected through the system created and the results from the questionnaire, it can not be concluded that all people that get more aware of their digital carbon footprint change their behavior on the Internet. It is believed that more research on making an application that helps the user change their behavior is needed, primarily focusing on the user interfaces.

This finding is strengthened by analyzing the feedback given in Questionnaire 2:

Feedback 7: I thought the idea was very interesting, and it was great to see the metrics that had been tracked based on my usage. I found it hard to understand what they meant, and so the dashboard overall didn't feel as impactful as it might have been if there was an onboarding explanation of what the metrics mean. Since I didn't have a foundational understanding either, the contextual comparison didn't feel meaningful either. The measurement I understood the most was "Bytes", the others were much less meaningful if at all. It was insightful to see where in the world the services I was using *were located in the world. It made me ponder the global reach of different online services..*

The primary resources of this thesis were used to develop the system's infrastructure needed to estimate the carbon footprint of individuals, and time spent on designing the user interfaces was, therefore, less prioritized. Thus, it can have reduced the impactfulness of the numbers and statistics presented to the user.

14.1.5 Discussing the Reseach Goal

By answering the research questions, the research goal *Characterize the effect visualizing an individual's digital carbon footprint has on the digital carbon footprint of the individual and the total carbon footprint of the Internet in the future* has been achieved.

This study demonstrates that visualizing a user's carbon footprint makes the user become more aware of their digital carbon footprint. Further, seven respondents answered agreeing to one or more of the questions in *Questionnaire 2, Part 2: Information about your Internet usage*, indicating that they changed their behavior on the Internet. Therefore, the effect of this experiment impacted the digital carbon footprint of some of the participants, but not all.

The study shows that visualizing the carbon footprint of individuals has less effect on participants that (1) have a low willingness to change their behavior to reduce their digital carbon footprint or (2) have a high estimated digital carbon footprint. Therefore, the effect of visualizing an individual's digital carbon footprint depends on the individual's Internet habits and willingness to change.

Because the effect of visualizing the climate footprint of individuals has a more negligible effect on people with an extensive digital climate footprint, the effect on the Internet's total carbon footprint is minimized.

When forecasting the future of the Internet's climate footprint, it is clear that behavioral change is needed. This study reveals that governmental legislation is necessary to reduce the ICT's climate footprint. It is believed that the only way to introduce these legislations is by societal change. The tipping point in social convention is 25% [27], hence the 38% of the participants that changed their behavior is enough to start a societal change enlightening the climate impact of the Internet. Consequently, visualizing the individuals' digital carbon footprint is considered superior for highlighting the current and influencing the future of the Internet's climate footprint.

14.2 Limitations

As with all research, this thesis also has its limitations. In this section, the different limitations of this study will be presented.

Calculations First of all, the calculation methods used to estimate the carbon footprint of the individuals can be said to be incomplete. Even though researchers often use CO_2e/GB or kWh/GB when estimating the carbon footprint of the Internet, research indicates that these figures can be unsuitable when estimating the carbon footprint of individuals [6]. As mentioned in Section 14.1, more complex mathematical formulas might be needed to calculate an estimation that is a better representation of the real world.

Limited to Web Browser Traffic Second, the system developed collected only the traffic sent through the web browser. This limited the ability of the system to estimate the complete carbon footprint of the participant. First of all, the system only tracked the size of the HTTP packets sent from the computer, not the IP packets. This resulted in the reported energy usage being systematically underestimated. Next, the system could not capture the data used on other devices. Streaming video is one of the most energy-demanding actions a user can perform on the Internet. Today, most people stream videos using applications on their Smart TVs or through technology like Chromecast and Apple TV. Not including the data usage from these devices can seriously underestimate the participants' carbon footprint. Statistics also show that around half of all Internet traffic happens through a mobile phone. Not including these numbers in the measures will also lead to underestimating a user's digital carbon footprint.

Excluding Land- and Water Footprint The system only estimated the carbon footprint of the participants without calculating the land- and water footprint. This limits the climate footprint of the participants only to include the carbon footprint and will therefore limit the end-users understanding of its digital climate footprint.

Sample Size 72 people signed up for the project, and of these 18 completed both questionnaires and installed the plugin. As the size of the total population of Internet users world wide was estimated to be roughly 4.7 billion in 2021 [38], the sample size for this thesis is small. When using a small sample size, research questions and hythesis can be quickly addressed and tested. However, a small sample size often leads to skew and not reliable or precise estimates [66]. For research, a confidence level of 95% with an accuracy range of 3% is usually chosen [109]. For the population of Internet users, a sample size of 1067 would be therefore needed to meet these criteria (using Cochran's formula [33]):

$$\frac{(1.96)^2 \cdot 0.5 \cdot 0.5}{0.3^2} \cdot 100 = 1067$$

As the sample size of 18 is significantly less than 1067, the results from the data analysis when answering both the research questions and research goal are directly affected. The limited number of participants can have affected the final results, and an incomplete picture can have been painted. Ideally, more participants should have answered both questionnaires and installed the plugin to ensure the results from the data generation methods replicated the real world.

Participants Most of the participants in this experiment were from Norway, and this can have affected the final results. Norway gets most of its energy from hydropower, which has a low carbon intensity. As the calculations in this thesis were done using the carbon intensity of the energy of the destination country of a network request, the numbers will be greater than the ones of Norway. This, in turn, can have made the estimations presented to the participants more impactful than if presented to users from a country where the carbon intensity of energy is higher than in Norway.

As self-selection sampling was used for this thesis, the participants might have had a strong feeling on the subject [109]. This can have affected the final results, as a higher percentage of the participants that signed up can have been interested in saving the environment and learning about the Internet's climate footprint than the average Internet user.

Time Frame The limited time frame of both the experiment and development of the final system can have impacted the execution of this project negatively. Ideally, the data collection through the plugin should have run for a longer time. Different events can affect the data usage of an individual considerably from day to day, and collecting data for more than two weeks would have reduced the impact of the individual variations in data usage. Having more time to develop and design the user interface of the final system could also have made the information presented to the end-user more impactful, and the time constrain therefore limited the full potential of the system.

14.3 Future work

Based on the work conducted in this thesis, several relevant focus points for future research have been identified.

Climate Footprint of the Internet As mentioned, the variations in the calculations of the Internet and its climate footprint are focus points for future work. More research should be executed in estimating the climate footprint of the Internet. Researchers should focus on laying pressure on governments to make ISPs and datacenters make their information about energy and data publicly available. More research should also be performed in the area of the land- and water footprint of the Internet to get a broader understanding of the climate footprint of the Internet. In addition, NFT's, increased usage of machine learning, higher-resolution TVs, and cryptocurrencies impact on the climate footprint of the Internet is a research point for the future.

Larger Sample Size As presented in Section 14.2, the sample size of this experiment can have affected the results of this thesis. In the future, researchers should focus on how to recruit more participants and reduce the dropout rate when developing systems to observe participants' carbon footprint.

Climate Footprint of Individuals The research executed in the field of ICT and its climate footprint tends to focus on the overall climate footprint of the Internet.

Researching formulas used to estimate the digital climate footprint of individuals should be prioritized for the future. This will make it easier to make applications that can change users' behavior on the Internet and be a crucial aid in reducing the carbon footprint of the ICT sector.

Regulations to Decrease the Data Usage of Organizations As mentioned, many of today's extensive ICT companies befit from consumers producing more and more data. In order to reduce the climate footprint of the Internet, a reduction in the data produced and consumed seems necessary. Research should be conducted to understand how regulations, such as a data tax, can be introduced to change the behavior of these IT companies and reduce the ICT sector's climate footprint.

A Framework for Developers to Develop Environmentally Friendly Applications The way systems are developed can significantly impact the environmental footprint of the user of such a system. In order to aid developers with developing environmentally friendly systems, future work should aim at creating a framework for developers that will be of aid when developing "earth first" applications.

Application that Captures More Internet Usage The system created in this thesis only estimated the carbon footprint from the web browser traffic. In the future, researchers should look into systems that can be created to estimate the user's carbon footprint across the user's many devices. This is believed to create a more accurate picture of the climate footprint of the user's Internet usage.

Impactful Application Finally, more research should be conducted to understand how the estimated carbon footprint of a user can be more impactful. During this research, most of the time was spent developing a system to estimate the user's carbon footprint. The infrastructure presented in this research can be used in the future. However, future work should analyze and look into how an application can significantly impact users' behavior on the Internet. The effect of the different features of the application created in this thesis can be of aid when developing a system for the future. Future work should prioritize changing the behavior of the groups that found the extension the least impactful, as presented in Chapter 13. The features ordered by popularity in Table 13.14 can be of good help to understand what features to prioritize, and what features that should not be included.

Machine Learning, Cryptocurrencies, and Non-Fungible Tokens (NFT) This thesis has not focused much on the fields of Machine Learning, Cryptocurrencies, and NFT's. However, the impact of and activity in all these fields can seriously increase the Internet's carbon footprint.

While the state-of-the-art machine learning algorithms are getting better, they also require more computational power when executed. This, in turn, increases the amount of energy used when running machine learning tasks, increasing the Internet's carbon footprint. As companies collect more and more data, there is a higher chance of them using machine learning to analyze the data. On one side, the use of computational heavy machine learning algorithms on large datasets can increase the Internet's carbon footprint. However, the results of the data analysis can also contribute to a reduction in the carbon footprint of the Internet if the results of the data analysis can be used to aid the companies or others in becoming more environmentally friendly.

Cryptocurrencies tend to become more and more popular and can significantly impact the ICT sector's carbon footprint going forward. Cryptocurrencies like Bitcoin require much power, both when mined and when making a transaction. Since 2012 the energy used by Bitcoin has had considerable growth and is estimated in 2022 to be 128.53 TWh [26].

Non-Fungible Token is another concept that can hurt the Internet's carbon footprint. With NFTs, individuals get ownership of digital assets, such as videos and images. Anyone can create assets to sell on the NFTs marketplaces. Seeing NFTs being sold for 69 million dollars [77] is believed to increase the number of people interested in starting their own NFT career. This, in turn, will result in more data being created, increasing the ICT sector's carbon footprint.

The impact of these technologies is a point to research for the future.

15 Conclusion

As part of this thesis, an application that estimates and gives feedback on an Internet user's carbon footprint was developed. Valuable insight into creating an application used to capture and visualize the digital carbon footprint has been provided. This thesis has highlighted the importance of focusing more on the ICT sector's climate footprint and showcased how an application can be created to make Internet users aware of their digital carbon footprint.

Through a literature review of the research conducted in estimating the Internet's carbon footprint, this thesis has emphasized some of the findings and challenges that exist within the research field. The results from the literature review were then used to plan out the technical requirements of an application able to estimate the carbon footprint of Internet users.

An interview with Gerry McGovern, the author of the book World Wide Waste, was conducted to explore his thoughts on visualizing the carbon footprint of Internet users. This thesis used a Google Chrome Extension to estimate the digital carbon footprint of an Internet user. A cloud-hosted API was used to estimate the carbon footprint, and the data was stored in a cloud database. The application was then developed, making it possible for the users to explore their own estimated digital carbon footprint worldwide and see their usage compared to the average of all the other participants.

A mixed-method approach was used to test the effect of the created application. At the beginning of the experiment, the participants answered a questionnaire to understand their Internet habits and knowledge about the Internet's climate footprint. The participants then installed the Chrome extension and used it for two weeks. During the first week, the application measured the participants' Internet usage in the background without the participants seeing their usage. In the second week, the participants were able to explore their usage and estimated carbon footprint. Finally, a questionnaire was answered by the participants to understand if the created application had helped them become more aware of their digital carbon footprint.

The answers from the questionnaire and the data collected through the applications were analyzed. The results from the questionnaire show that the created application made 83% of the participants more aware of their digital carbon footprint. However, the research also shows that people who become aware of their digital carbon footprint do not necessarily change their behavior on the Internet, as 11 out of 18 did not change their behavior. Nevertheless, the 38% that did change their behavior is enough for a societal change towards a more sustainable Internet. Thus, this research contributes to making the Internet a mainstream climate concern.

Bibliography

[1] France 24. Chill your netflix habit, climate experts say - france 24. https://www.france24.com/en/20191028-chill-your-netflix-habit-climate-experts-say, 2019. (Accessed on 12/14/2021).

[2] European Environment Agency. Energy and climate change — european environment agency. https://www.eea.europa.eu/signals/signals-2017/articles/energy-and-climate-change, 08 2017. (Accessed on 02/18/2022).

[3] IEA International Energy Agency. The carbon footprint of streaming video: fact-checking the headlines. https://www.iea.org/commentaries/the-carbon-footprint-of-streaming-video-fact-checking-the-headlines, 2020. (Accessed on 12/14/2021).

[4] Blake Alcott. Jevons' paradox. *Ecological economics*, 54(1):9–21, 2005.

[5] Anders SG Andrae. Projecting the chiaroscuro of the electricity use of communication and computing from 2018 to 2030. *Preprint*, pages 1–23, 2019.

[6] Anders SG Andrae. New perspectives on internet electricity use in 2030. *Engineering and Applied Science Letter*, 3(2):19–31, 2020.

[7] Anders SG Andrae and Tomas Edler. On global electricity usage of communication technology: trends to 2030. *Challenges*, 6(1):117–157, 2015.

[8] Apple. Safari web extensions | apple developer documentation. https://developer.apple.com/documentation/safariservices/safari_web_extensions. (Accessed on 01/04/2022).

[9] Apple. Use focus on your iphone, ipad, or ipod touch - apple support. https://support.apple.com/en-us/HT212608, September 2021. (Accessed on 01/06/2022).

[10] HTTP Archive. Page weight | 2021 | the web almanac by http archive. https://almanac.httparchive.org/en/2021/page-weight#page-weight-by-the-numbers. (Accessed on 12/09/2021).

[11] M.F. Arlitt and C.L. Williamson. Internet web servers: workload characterization and performance implications. *IEEE/ACM Transactions on Networking*, 5(5):631–645, 1997.

[12] Joshua Aslan, Kieren Mayers, Jonathan G. Koomey, and Chris France. Electricity intensity of internet data transmission: Untangling the estimates. *Journal of Industrial Ecology*, 22(4):785–798, 2018.

[13] V. asson Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, J.B.R. Matthews E. Lonnoy, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.). Climate change 2021: The physical science basis. contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change. *Cambridge University Press. In Press.*, 2021.

[14] Auth0. Auth0: Secure access for everyone. but not just anyone. `https://auth0.com/`, 2022. (Accessed on 01/11/2022).

[15] AWS. What is continuous integration? – amazon web services. `https://aws.amazon.com/devops/continuous-integration/`. (Accessed on 01/14/2022).

[16] Umberto Bacchi. How cat videos could cause a 'climate change nightmare' | reuters. `https://www.reuters.com/article/us-georgia-tech-climatechange-feature-tr-idUSKBN20C1A7`, 2020. (Accessed on 12/14/2021).

[17] Jayant Baliga, Robert Ayre, Kerry Hinton, Wayne V Sorin, and Rodney S Tucker. Energy consumption in optical ip networks. *Journal of Lightwave Technology*, 27(13):2391–2403, 2009.

[18] Thomas Barnett, Shruti Jain, Usha Andra, and Taru Khurana. Cisco visual networking index (vni) complete forecast update, 2017–2022. *Americas/EMEAR Cisco Knowledge Network (CKN) Presentation*, 2018.

[19] John Bartholomew. Is there a purpose for using pull requests on my own repo if i am the only developer? `https://softwareengineering.stackexchange.com/a/178489`, 2012. (Accessed on 01/20/2022).

[20] Kent Beck. *Test-driven development: by example*. Addison-Wesley Professional, 2003.

[21] India Berry. Top 10 countries with the most data centres | data centre magazine. `https://datacentremagazine.com/top10/top-10-countries-most-data-centres`, 2021. (Accessed on 02/11/2022).

[22] Alex Berson. *Client/server architecture*. McGraw-Hill, Inc., 1996.

[23] Tova Billstein, Anna Björklund, and Tomas Rydberg. Life cycle assessment of network traffic: A review of challenges and possible solutions. *Sustainability*, 13(20):11155, 2021.

[24] Timm Böttger, Ghida Ibrahim, and Ben Vallis. How the internet reacted to covid-19: A perspective from facebook's edge network. In *Proceedings of the ACM Internet Measurement Conference*, pages 34–41, 2020.

[25] Victor R Basili1 Gianluigi Caldiera and H Dieter Rombach. The goal question metric approach. *Encyclopedia of software engineering*, pages 528–532, 1994.

[26] Cambridge. Cambridge bitcoin electricity consumption index (cbeci). `https://ccaf.io/cbeci/index`, 2022. (Accessed on 02/02/2022).

[27] Damon Centola, Joshua Becker, Devon Brackbill, and Andrea Baronchelli. Experimental evidence for tipping points in social convention. *Science*, 360(6393):1116–1119, 2018.

[28] Chrome. Content scripts - chrome developers. `https://developer.chrome.com/docs/extensions/mv3/content_scripts/`, 2021. (Accessed on 01/14/2022).

[29] Chromium. The chromium projects. `https://www.chromium.org/`. (Accessed on 01/04/2022).

[30] U Cisco. Cisco annual internet report (2018–2023) white paper, 2020.

[31] cletus and beppe9000. How to get started building a web browser? [closed]. `https://stackoverflow.com/q/598841/11722925`. (Accessed on 01/04/2022).

[32] Google Cloud. App engine application platform | google cloud. `https://cloud.google.com/appengine`, 2022. (Accessed on 01/17/2022).

[33] William Gemmell Cochran. *Sampling Techniques: 3d Ed.* Wiley, 1977.

[34] MDN Contributors. What are extensions? - mozilla | mdn. `https://developer.mozilla.org/en-US/docs/Mozilla/Add-ons/WebExtensions/What_are_WebExtensions`, 2021. (Accessed on 01/04/2022).

[35] Vlad C Coroama, Lorenz M Hilty, Ernst Heiri, and Frank M Horn. The direct energy demand of internet data flows. *Journal of Industrial Ecology*, 17(5):680–688, 2013.

[36] David Costenaro and Anthony Duer. The megawatts behind your megabytes: going from data-center to desktop. *Proceedings of the 2012 ACEEE Summer Study on Energy Efficiency in Buildings, ACEEE, Washington*, pages 13–65, 2012.

[37] Jeannette Cwienk. Is netflix bad for the environment? how streaming video contributes to climate change. `https://www.dw.com/en/is-netflix-bad-for-the-environment-how-streaming-video-contributes-to-climate-change/a-49556716`, 2019. (Accessed on 12/14/2021).

[38] DataReportal. Digital 2021 october global statshot report — datareportal – global digital insights. `https://datareportal.com/reports/digital-2021-october-global-statshot`, 2021. (Accessed on 01/04/2022).

[39] Wholegrain Digital. How does it work? - website carbon calculator. `https://www.websitecarbon.com/how-does-it-work/`, 2019. (Accessed on 01/07/2022).

[40] Rich Edwards. Sql versus nosql: Pros and cons | datastax. `https://www.datastax.com/blog/sql-vs-nosql-pros-cons`, 2021. (Accessed on 01/11/2022).

[41] Maxime Efoui-Hess. Climate crisis: The unsustainable use of online video. *The Shift Project: Paris, France*, 2019.

[42] Ingrid Ekeberg. Ny rapport: Global videostrømming står for like store utslipp som spania | dn. `https://www.dn.no/klima/ny-rapport-global-videostromming-star-for-like-store-utslipp-som-spania/2-1-636528`, 2019. (Accessed on 12/14/2021).

[43] Global Energy. Co2 status report. *IEA (International Energy Agency): Paris, France*, 2019.

[44] Narayanan A. Englehardt S. Online tracking: A 1-million-site measurement and analysis. 2016.

[45] esbuild. esbuild - an extremely fast javascript bundler. `https://esbuild.github.io/`, 2022. (Accessed on 01/11/2022).

[46] ESLint. About - eslint - pluggable javascript linter. `https://eslint.org/docs/about/`, 2022. (Accessed on 01/12/2022).

[47] Express. Using express middleware. `https://expressjs.com/en/guide/using-middleware.html`, 2021. (Accessed on 01/17/2022).

[48] Facebook. facebook/react: A declarative, efficient, and flexible javascript library for building user interfaces. `https://github.com/facebook/react`, 2021. (Accessed on 01/12/2022).

[49] Anja Feldmann, Oliver Gasser, Franziska Lichtblau, Eric Pujol, Igmar Poese, Christoph Dietzel, Daniel Wagner, Matthias Wichtlhuber, Juan Tapiador, Narseo Vallina-Rodriguez, et al. Implications of the covid-19 pandemic on the internet traffic. In *Broadband Coverage in Germany; 15th ITG-Symposium*, pages 1–5. VDE, 2021.

[50] Firebase. Cloud firestore | firebase documentation. `https://firebase.google.com/docs/firestore`, 2022. (Accessed on 01/11/2022).

[51] Firebase. Firebase authentication | simple, free multi-platform sign-in. `https://firebase.google.com/products/auth`, 2022. (Accessed on 01/11/2022).

[52] Firebase. Firebase hosting | fast and secure web hosting. `https://firebase.google.com/products/hosting?gclid=Cj0KCQiAip-PBhDVARIsAPP2xc1gySMGreqpEASbJ_70wR8rMu-IV0L_d_44kG0mqEk4YMH_RWVC_tkaAru7EALw_wcB&gclsrc=aw.ds`, 2022. (Accessed on 01/19/2022).

[53] Firebase. Introduction to firebase local emulator suite | firebase documentation. `https://firebase.google.com/docs/emulator-suite`, 2022. (Accessed on 01/14/2022).

[54] Kenneth Flamm. Measuring moore's law: evidence from price, cost, and quality indexes. Technical report, National Bureau of Economic Research, 2018.

[55] Carbon Footprint. Country specific electricity grid greenhouse gas emission factors. *Carbon Footprint: Hampshire, UK*, 2019.

[56] Charlotte Freitag, Mike Berners-Lee, Kelly Widdicks, Bran Knowles, Gordon Blair, and Adrian Friday. The climate impact of ict: A review of estimates, trends and regulations. *arXiv preprint arXiv:2102.02622*, 2021.

[57] GCFGlocal. Internet safety: Understanding browser tracking. `https://edu.gcfglobal.org/en/internetsafety/understanding-browser-tracking/1/`, 2021. (Accessed on 12/13/2021).

[58] Geeksforgeeks. Singleton design pattern | implementation - geeksforgeeks. `https://www.geeksforgeeks.org/singleton-design-pattern/`, 2020. (Accessed on 01/17/2022).

[59] J Geist. Did the shift project really overestimate the carbon footprint of online video? our analysis of the iea and carbonbrief articles. 2020.

[60] Github. About. `https://github.com/about`, 2022. (Accessed on 01/19/2022).

[61] Martin Glinz. On non-functional requirements. In *15th IEEE international requirements engineering conference (RE 2007)*, pages 21–26. IEEE, 2007.

[62] Google. Api reference - chrome developers. `https://developer.chrome.com/docs/extensions/reference/`. (Accessed on 01/04/2022).

[63] Google. Digital wellbeing tools | google. `https://wellbeing.google/for-everyone/`. (Accessed on 01/06/2022).

[64] Google. Sustainability commitments for our future | google sustainability. `https://sustainability.google/commitments/`. (Accessed on 02/01/2022).

[65] Google. Overcoming technology challenges | digital wellbeing | google. `https://wellbeing.google/our-commitment/`, 2022. (Accessed on 02/11/2022).

[66] Allan Hackshaw. Small studies: strengths and limitations, 2008.

[67] Ola Hansén and Nina Schmieder. Wwfs klimatbarometer: Allt fler väljer bort flyg och kött – och kvinnorna går före - världsnaturfonden wwf. `https://www.wwf.se/pressmeddelande/wwfs-klimatbarometer-allt-fler-valjer-bort-flyg-och-kott-och-kvinnorna-gar-fore-3241404/`, 2019. (Accessed on 02/04/2022).

[68] Lorenz M Hilty and Bernard Aebischer. *ICT innovations for sustainability*, volume 310. Springer, 2015.

[69] Ralph Hintemann and Simon Hinterholzer. Energy consumption of data centers worldwide. In *The 6th International Conference on ICT for Sustainability (ICT4S). Lappeenranta*, 2019.

[70] IBM. Http requests - ibm documentation. `https://www.ibm.com/docs/en/cics-ts/5.3?topic=protocol-http-requests`, 2021. (Accessed on 12/13/2021).

[71] IBM. Http responses - ibm documentation. `https://www.ibm.com/docs/en/cics-ts/5.3?topic=protocol-http-responses`, 20211. (Accessed on 12/13/2021).

[72] René Itten, Roland Hischier, Anders SG Andrae, Jan CT Bieser, Livia Cabernard, Annemarie Falke, Hugues Ferreboeuf, Lorenz M Hilty, Regula L Keller, Etienne Lees-Perasso, et al. Digital transformation—life cycle assessment of digital services, multifunctional devices and cloud computing. *The International Journal of Life Cycle Assessment*, 25(10):2093–2098, 2020.

[73] JavaScript. Javascript.com. `https://www.javascript.com/`. (Accessed on 12/10/2021).

[74] Jest. Jest · delightful javascript testing. `https://jestjs.io/`, 2022. (Accessed on 01/17/2022).

[75] William Stanley Jevons. *The coal question; an inquiry concerning the progress of the nation and the probable exhaustion of our coal-mines.* 1866.

[76] Ankur Joshi, Saket Kale, Satish Chandel, and D Kumar Pal. Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4):396, 2015.

[77] Jacob Kastrenakes. Beeple sold an nft for $69 million - the verge. `https://www.theverge.com/2021/3/11/22325054/beeple-christies-nft-sale-cost-everydays-69-million`, Mars 2021. (Accessed on 02/02/2022).

[78] Keycloak. Keycloak. `https://www.keycloak.org/`, 2021. (Accessed on 01/11/2022).

[79] Jonathan Koomey, Huimin Chong, Woonsien Loh, Bruce Nordman, and Michele Blazek. Network electricity use associated with wireless personal digital assistants. *Journal of infrastructure systems*, 10(3):131–137, 2004.

[80] Jonathan Koomey and Eric Masanet. Does not compute: Avoiding pitfalls assessing the internet's energy and carbon impacts. *Joule*, 5(7):1625–1628, 2021.

[81] Martijn Koot and Fons Wijnhoven. Usage impact on data center electricity needs: A system dynamic forecasting model. *Applied Energy*, 291:116798, 2021.

[82] Louise Krug, Mark Shackleton, and Fabrice Saffre. Understanding the environmental costs of fixed line networking. In *Proceedings of the 5th international conference on Future energy systems*, pages 87–95, 2014.

[83] Arwa Mahdawi. The real problem with your netflix addiction? the carbon emissions | | the guardian. `https://www.theguardian.com/commentisfree/2020/feb/12/real-problem-netflix-addiction-arbon-emissions`, 2020. (Accessed on 12/14/2021).

[84] Jens Malmodin. The power consumption of mobile and fixed network data services-the case of streaming video and downloading large files. *Electronics Goes Green 2020*, 2020.

[85] Jens Malmodin and Dag Lundén. The energy and carbon footprint of the ict and e&m sector in sweden 1990-2015 and beyond. *ICT for Sustainability*, 2016.

[86] Jens Malmodin, Dag Lundén, Åsa Moberg, Greger Andersson, and Mikael Nilsson. Life cycle assessment of ict: Carbon footprint and operational electricity use from the operator, national, and subscriber perspective in sweden. *Journal of Industrial Ecology*, 18(6):829–845, 2014.

[87] Laura U Marks, Stephen Makonin, Alejandro Rodriguez-Silva, and Radek Przedpełski. Tackling the carbon footprint of streaming media. `https://www.sshrc-crsh.gc.ca/society-societe/community-communite/ifca-iac/evidence_briefs-donnees_probantes/earth_carrying_capacity-capacite_limite_terre/marks_makonin_przedpelski_rodriguez-silva-eng.aspx`.

[88] E Masanet and Lei N. How much energy do data centers really use? | aspen global change institute. 2020. (Accessed on 12/14/2021).

[89] Ian McDonough. Making video more sustainable. `https://www.forbes.com/sites/forbestechcouncil/2021/08/31/making-video-more-sustainable/`, 2021. (Accessed on 12/14/2021).

[90] Gerry McGovern. Home - gerry mcgovern. `https://gerrymcgovern.com/`. (Accessed on 01/05/2022).

[91] Martin McKeay. Akamai blog | the building wave of internet traffic. `https://www.akamai.com/blog/security/the-building-wave-of-internet-traffic`, April 2020. (Accessed on 01/04/2022).

[92] Patrick E McKnight and Julius Najab. Mann-whitney u test. *The Corsini encyclopedia of psychology*, pages 1–1, 2010.

[93] Microsoft. Azure sustainability—sustainable technologies | microsoft azure. `https://azure.microsoft.com/en-us/global-infrastructure/sustainability/`, 2018. (Accessed on 02/01/2022).

[94] Paul Monk. To sql or not to sql. https://capgemini.github.io/design/sql-vs-nosql/, 2017. (Accessed on 01/11/2022).

[95] F Montevecchi, Therese Stickler, R Hintemann, and S Hinterholzer. Energy-efficient cloud computing technologies and policies for an eco-friendly cloud market. *European Commission, final study report, Vienna, report commissioned by the Directorate-General for Communications Networks, Content and Technology*, 2020.

[96] Camilo Mora, Daniele Spirandelli, Erik C Franklin, John Lynham, Michael B Kantar, Wendy Miles, Charlotte Z Smith, Kelle Freel, Jade Moy, Leo V Louis, et al. Broad threat to humanity from cumulative climate hazards intensified by greenhouse gas emissions. *Nature Climate Change*, 8(12):1062–1071, 2018.

[97] Frank Moraes. Html for beginners the easy way: Start learning html & css today ». https://html.com/, 2021. (Accessed on 12/13/2021).

[98] Mozilla. Evolution of http - http | mdn. https://developer.mozilla.org/en-US/docs/Web/HTTP/Basics_of_HTTP/Evolution_of_HTTP, 2021. (Accessed on 12/13/2021).

[99] Mozilla. Fetch api - web apis | mdn. https://developer.mozilla.org/en-US/docs/Web/API/Fetch_API, 2021. (Accessed on 01/14/2022).

[100] Mozilla. Http | mdn. https://developer.mozilla.org/en-US/docs/Web/HTTP, 2021. (Accessed on 12/13/2021).

[101] Mozilla. Http headers - http | mdn. https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers, 2021. (Accessed on 12/13/2021).

[102] Mozilla. Http request methods - http | mdn. https://developer.mozilla.org/en-US/docs/Web/HTTP/Methods, 2021. (Accessed on 12/13/2021).

[103] Mozilla. Http response status codes - http | mdn. https://developer.mozilla.org/en-US/docs/Web/HTTP/Status, 2021. (Accessed on 12/13/2021).

[104] Mozilla. Mvc - mdn web docs glossary: Definitions of web-related terms | mdn. https://developer.mozilla.org/en-US/docs/Glossary/MVC, 2021. (Accessed on 01/17/2022).

[105] Mozilla. Xmlhttprequest - web apis | mdn. https://developer.mozilla.org/en-US/docs/Web/API/XMLHttpRequest, 2021. (Accessed on 01/14/2022).

[106] Subramanian Senthilkannan Muthu. *Assessing the environmental impact of textiles and the clothing supply chain*. Woodhead publishing, 2020.

[107] Tomoko Nemoto and David Beglar. Likert-scale questionnaires. In *JALT 2013 conference proceedings*, pages 1–8, 2014.

[108] Sevanti Ninan. The carbon cost of digitization is an under-reported story - telegraph india. https://www.telegraphindia.com/opinion/the-carbon-cost-of-digitization-is-an-under-reported-story/cid/1839969, 2021. (Accessed on 12/14/2021).

[109] Briony J Oates. *Researching information systems and computing*. Sage, 2005.

[110] Renee Obringer, Benjamin Rachunok, Debora Maia-Silva, Maryam Arbabzadeh, Roshanak Nateghi, and Kaveh Madani. The overlooked environmental footprint of increasing internet use. *Resources, Conservation and Recycling*, 167:105389, 2021.

[111] Ofcom. Online nation - ofcom. https://www.ofcom.org.uk/research-and-data/internet-and-on-demand-research/online-nation, 2021. (Accessed on 01/05/2022).

[112] Ookla. Despite all odds, global internet speeds continue impressive increase. https://www.speedtest.net/insights/blog/world-internet-speeds-july-2021/, 2021. (Accessed on 01/07/2022).

[113] Jeff Overton. Fact sheet | the growth in greenhouse gas emissions from commercial aviation (2019) | white papers | eesi. https://www.eesi.org/papers/view/fact-sheet-the-growth-in-greenhouse-gas-emissions-from-commercial-aviation, 2019. (Accessed on 01/28/2022).

[114] Parcel. Parcel. https://parceljs.org/, 2022. (Accessed on 01/11/2022).

[115] Mike Peterson. Snap, facebook using loopholes to bypass app tracking transparency | appleinsider. https://appleinsider.com/articles/21/12/10/snap-facebook-using-loopholes-to-bypass-app-tracking-transparency?utm_medium=rss, 2021. (Accessed on 12/13/2021).

[116] Lorenzo Pichilli. pichillilorenzo/flutter_browser_app: A full-featured mobile browser app (such as the google chrome mobile browser) created using flutter and the features offered by the flutter_inappwebview plugin. https://github.com/pichillilorenzo/flutter_browser_app. (Accessed on 01/04/2022).

[117] Lorenzo Pichilli. pichillilorenzo/flutter_inappwebview: A flutter plugin that allows you to add an inline webview, to use a headless webview, and to open an in-app browser window. https://github.com/pichillilorenzo/flutter_inappwebview. (Accessed on 01/04/2022).

[118] pnpm. Motivation | pnpm. https://pnpm.io/motivation, 2022. (Accessed on 01/19/2022).

[119] PostgreSQL. Postgresql: The world's most advanced open source database. https://www.postgresql.org/. (Accessed on 01/11/2022).

[120] Preact. Preact | preact: Fast 3kb react alternative with the same es6 api. components & virtual dom. `https://preactjs.com/`, 2022. (Accessed on 01/12/2022).

[121] Max Rekhopf. What is a kanban board? | atlassian. `https://www.atlassian.com/agile/kanban/boards`. (Accessed on 01/20/2022).

[122] Interactive Results. Interactive results. `https://krausest.github.io/js-framework-benchmark/2021/table_chrome_93.0.4577.63.html`, 2021. (Accessed on 01/11/2022).

[123] RhodeCode. Rhodecode › version control systems popularity in 2016. `https://rhodecode.com/insights/version-control-systems-2016`, 2016. (Accessed on 01/19/2022).

[124] D Ruiz, G San Miguel, J Rojo, JG Teriús-Padrón, E Gaeta, MT Arredondo, JF Hernández, and J Pérez. Life cycle inventory and carbon footprint assessment of wireless ict networks for six demographic areas. *Resources, Conservation and Recycling*, 176:105951, 2022.

[125] Nayanamana Samarasinghe and Mohammad Mannan. Towards a global perspective on web tracking. *Computers & Security*, 87:101569, 2019.

[126] Dan Schien and Chris Preist. Approaches to energy intensity of the internet. *IEEE Communications Magazine*, 52(11):130–137, 2014.

[127] Arman Shehabi, Ben Walker, and Eric Masanet. The energy and greenhouse-gas implications of internet video streaming in the united states. *Environmental Research Letters*, 9(5):054007, 2014.

[128] Navid Siami and Ralph A Winter. Jevons' paradox revisited: Implications for climate change. *Economics Letters*, page 109955, 2021.

[129] Statista. Chart: How many websites are there? `https://www.statista.com/chart/19058/number-of-websites-online/`, 2021. (Accessed on 12/09/2021).

[130] Statscounter. Desktop vs mobile vs tablet market share worldwide | statcounter global stats. `https://gs.statcounter.com/platform-market-share/desktop-mobile-tablet`, 2021. (Accessed on 01/04/2022).

[131] Hilda Stenberg and Anders Bjäerlestam. Svt play. `https://youtu.be/XoOPB5i_b4Y?t=3286`. (Accessed on 01/06/2022).

[132] Dan Sullivan. *NoSQL for mere mortals*. Addison-Wesley Professional, 2015.

[133] Tailwind. Tailwind css - rapidly build modern websites without ever leaving your html. `https://tailwindcss.com/`, 2022. (Accessed on 01/17/2022).

[134] Cody Taylor and Jonathan Koomey. Estimating energy use and greenhouse gas emissions of internet advertising. *Network*, 2008.

[135] Carbon Trust. Carbon impact of video streaming. https://www.carbontrust.com/resources/carbon-impact-of-video-streaming, 2021. (Accessed on 12/14/2021).

[136] Nina Tuv and Emmie S. Foss. Mer tid til strømming og spill i koronaåret - ssb. https://www.ssb.no/kultur-og-fritid/artikler-og-publikasjoner/mer-tid-til-stromming-og-spill-i-koronaaret, 04 2021. (Accessed on 01/05/2022).

[137] TypeScript. Typescript: Javascript with syntax for types. https://www.typescriptlang.org/, 2022. (Accessed on 01/12/2022).

[138] Dr Ullah. Can i use a paired samples t-test for 5-point likert scale data?, 09 2015.

[139] ITU-T International Telecommunication Union. Greenhouse gas emissions trajectories for the information and communication technology sector compatible with the unfccc paris agreement. https://www.itu.int/rec/T-REC-L.1470-202001-I/en, 2020. (Accessed on 12/14/2021).

[140] Vercel. Next.js by vercel - the react framework. https://nextjs.org/, 2022. (Accessed on 01/19/2022).

[141] Vue. Introduction | vue.js. https://v3.vuejs.org/guide/introduction.html, 2022. (Accessed on 01/11/2022).

[142] W3. Cascading style sheets. https://www.w3.org/Style/CSS/Overview.en.html. (Accessed on 12/12/2021).

[143] W3Schools. Browser statistics. https://www.w3schools.com/browsers/, 2021. (Accessed on 01/11/2022).

[144] Bill Wake. Invest in good stories, and smart tasks - xp123. https://xp123.com/articles/invest-in-good-stories-and-smart-tasks/, 2003. (Accessed on 01/20/2022).

[145] Christine Wamsler and Ebba Brink. Mindsets for sustainability: Exploring the link between mindfulness and sustainable climate adaptation. *Ecological Economics*, 151:55–61, 2018.

[146] Samuel Webb. Grocer introduces receipt with carbon footprint of weekly shop | the independent. https://www-independent-co-uk.cdn.ampproject.org/c/s/www.independent.co.uk/climate-change/news/oda-red-meat-carbon-footprint-b1988091.html?amp, 2022. (Accessed on 02/05/2022).

[147] Christopher L Weber, Jonathan G Koomey, and H Scott Matthews. The energy and climate change implications of different music delivery methods. *Journal of Industrial Ecology*, 14(5):754–769, 2010.

[148] Webpack. webpack. `https://webpack.js.org/`, 2021. (Accessed on 12/13/2021).

[149] Daniel R Williams and Yinshan Tang. Methodology to model the energy and greenhouse gas emissions of electronic software distributions. *Environmental science & technology*, 46(2):1087–1095, 2012.

[150] Claes Wohlin. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th international conference on evaluation and assessment in software engineering*, pages 1–10, 2014.

[151] Robert F Woolson. Wilcoxon signed-rank test. *Wiley encyclopedia of clinical trials*, pages 1–3, 2007.

[152] Evan You. vuejs/vue: Vue.js is a progressive, incrementally-adoptable javascript framework for building ui on the web. `https://github.com/vuejs/vue`, 2022. (Accessed on 01/11/2022).

[153] Bartłomiej Żyliński. What keycloak is and what it does? - dzone security. `https://dzone.com/articles/what-is-keycloak-and-when-it-may-help-you`, 2021. (Accessed on 01/11/2022).

Appendices

1 Interview Guide

This appendix shows the interview guide used during the Interview with Gerry McGovern. The Interview was structured as a semi-structured interview.

Introduction
Tell me a bit about yourself
Your work with the Internet and it's climate impact

The environment and the Internet
When did you become aware of the climate footprint?
Are people aware of the Internet's climate footprint?
How do you think that we can reduce the Internet's climate footprint?
Who's responsible for leading the change towards a less polluting internet?
Why do you think there is so little focus on the internet and its climate consequences?

The Plugin
What do you think of the idea of having a plugin that measures your digital climate footprint?
What kind of information to give users to make them change behavior?
How to present the information that's needed to change the behavior?
How long time do you think people need before they change their behavior on the internet?
Can people be nudged to pollute less?
What do you think will be the biggest challenge with this plugin?

2 Security Rules for Firestore

As privacy was of concern for this thesis, security rules was in place to limit the access to the data stored in Firestore. In this appendix I will present the security rules and how they worked.

```
1  rules_version = '2';
2  service cloud.firestore {
3    match /databases/{database}/documents {
4      match /usage/{anyDocument} {
5        allow read: if request.auth != null && (request.auth.uid ==
             resource.data.userId || !exists(request.path))
6      }
7      match /countries/{anyDocument} {
8        allow read: if request.auth != null && (request.auth.uid ==
             resource.data.userId || !exists(request.path))
9      }
10     match /hosts/{anyDocument} {
11       allow read: if request.auth != null && (request.auth.uid ==
             resource.data.userId || !exists(request.path))
12     }
13     match /hostToCountry/{anyDocument} {
14       allow read: if request.auth != null && (request.auth.uid ==
             resource.data.userId || !exists(request.path))
15     }
16     match /users/{userId}/{documents=**} {
17       allow read: if request.auth != null && request.auth.uid == userId
18     }
19     match /{document=**} {
20       allow read, write: if false;
21     }
22   }
23 }
```

The rules works by matching different paths of the document database. They work in a top-down matter, where the first rule overrules the last rule. For the collections *usage*, *countries*, *hosts*, *hostToCountry* and *users*. As I wanted to use the realtime functionality of Firestore for these collections, read rules was applied. The rule defined access to the users that was logged in and that was the owner of the document. If the path didn't exist, they were also granted access to read, to make sure that the realtime update wouldn't break when trying to read documents that didn't yet exists. For the rest of the CRUD operations on the rest of the documents, no access was given for the client, as shown on line 19-21. Access to these collections had to go through the API with its own security rules in place.

3 CO_2 equivalents (kg) per kWh per country

The following JSON object was created based on the article *Country specific electricity grid greenhouse gas emission factors, reporting the country specific electricity grid greenhouse gas emission factors for 2019* [55].


```
{
  "ZA": {
    "country": "South Africa",
    "CO2perkWh": 0.928,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "CN": {
    "country": "China",
    "CO2perkWh": 0.555,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "HK": {
    "country": "Hong Kong",
    "CO2perkWh": 0.81,
    "source": "Hong Kong Electric Company (2019) or CLP Group (2019)
       These two companies supply different areas of HK so check which
       one you need.",
    "year": 2019
  },
  "IN": {
    "country": "India",
    "CO2perkWh": 0.708,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "ID": {
    "country": "Indonesia",
    "CO2perkWh": 0.761,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "JP": {
    "country": "Japan",
    "CO2perkWh": 0.506,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "KR": {
    "country": "Korea",
    "CO2perkWh": 0.5,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "SG": {
    "country": "Singapore",
```

```
    "CO2perkWh": 0.4188,
    "source": "Singapore Energy Market Authority (EMA)",
    "year": 2018
  },
  "TH": {
    "country": "Thailand",
    "CO2perkWh": 0.497,
    "source": "Energy Policy and Planning Office (EPPO) Thai Government Ministry of Energy",
    "year": 2019
  },
  "AU": {
    "country": "Australia",
    "CO2perkWh": 0.88,
    "source": "Australian Government",
    "year": 2018
  },
  "NZ": {
    "country": "New Zealand",
    "CO2perkWh": 0.1051,
    "source": "Ministry for the Environment https://www.mfe.govt.nz/node/18670/",
    "year": null
  },
  "SA": {
    "country": "Saudi Arabia",
    "CO2perkWh": 0.732,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "TR": {
    "country": "Turkey",
    "CO2perkWh": 0.481,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "AE": {
    "country": "United Arab Emirates",
    "CO2perkWh": 0.4258,
    "source": "Dubai Electricity & Water Authority (sustainability report 2018)",
    "year": 2018
  },
  "CA": {
    "country": "Canada",
    "CO2perkWh": 0.13,
    "source": "UN Framework Convention on Climate Change",
    "year": null
  },
  "MX": {
    "country": "Mexico",
    "CO2perkWh": 0.449,
    "source": "Climate Transparency (2019 Report)",
```

```
    "year": 2018
  },
  "US": {
    "country": "United States",
    "CO2perkWh": 0.45322,
    "source": "US Env Protection Agency (EPA) eGrid",
    "year": 2018
  },
  "AR": {
    "country": "Argentina",
    "CO2perkWh": 0.313,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "BR": {
    "country": "Brazil",
    "CO2perkWh": 0.074,
    "source": "Climate Transparency (2019 Report)",
    "year": 2018
  },
  "AT": {
    "country": "Austria",
    "CO2perkWh": 0.13286,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "BE": {
    "country": "Belgium",
    "CO2perkWh": 0.15313,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "BG": {
    "country": "Bulgaria",
    "CO2perkWh": 0.43737,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "HR": {
    "country": "Croatia",
    "CO2perkWh": 0.27315,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "CY": {
    "country": "Cyprus",
    "CO2perkWh": 0.67729,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "CZ": {
    "country": "Czech Republic",
    "CO2perkWh": 0.54465,
```

```
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "DK": {
    "country": "Denmark",
    "CO2perkWh": 0.15444,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "EE": {
    "country": "Estonia",
    "CO2perkWh": 0.72328,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "FI": {
    "country": "Finland",
    "CO2perkWh": 0.13622,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "FR": {
    "country": "France",
    "CO2perkWh": 0.03895,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "DE": {
    "country": "Germany",
    "CO2perkWh": 0.37862,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "GR": {
    "country": "Greece",
    "CO2perkWh": 0.54901,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "HU": {
    "country": "Hungary",
    "CO2perkWh": 0.25298,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "IS": {
    "country": "Iceland",
    "CO2perkWh": 0.00011,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "IE": {
    "country": "Ireland",
```

```
    "CO2perkWh": 0.34804,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "IT": {
    "country": "Italy",
    "CO2perkWh": 0.33854,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "LV": {
    "country": "Latvia",
    "CO2perkWh": 0.30333,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "LT": {
    "country": "Lithuania",
    "CO2perkWh": 0.14913,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "LU": {
    "country": "Luxembourg",
    "CO2perkWh": 0.13939,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "MT": {
    "country": "Malta",
    "CO2perkWh": 0.3706,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "NL": {
    "country": "Netherlands",
    "CO2perkWh": 0.45207,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "NO": {
    "country": "Norway",
    "CO2perkWh": 0.01118,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "PL": {
    "country": "Poland",
    "CO2perkWh": 0.79107,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "PT": {
```

```
    "country": "Portugal",
    "CO2perkWh": 0.25255,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "RO": {
    "country": "Romania",
    "CO2perkWh": 0.31011,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "RU": {
    "country": "Russian Federation",
    "CO2perkWh": 0.325,
    "source": "Climate Transparency (2019 Report)",
    "year": 2019
  },
  "RS": {
    "country": "Serbia",
    "CO2perkWh": 0.76253,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "SK": {
    "country": "Slovakia",
    "CO2perkWh": 0.1511,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "SI": {
    "country": "Slovenia",
    "CO2perkWh": 0.24385,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "ES": {
    "country": "Spain",
    "CO2perkWh": 0.22026,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "SE": {
    "country": "Sweden",
    "CO2perkWh": 0.01189,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
  "CH": {
    "country": "Switzerland",
    "CO2perkWh": 0.01182,
    "source": "Association of Issuing Bodies (AIB)",
    "year": 2019
  },
```

```
  "GB": {
    "country": "United Kingdom",
    "CO2perkWh": 0.2531897,
    "source": "Production: UK Govt âĂŞ Defra/BEIS Residual: Association of Issuing Bodies (AIB)",
    "year": undefined
  }
}
```

4 User Testing Guide

During the development of the application two user tests were conducted. In this appendix, the user test is guide shown.

1. Go to the given URL and install the plugin. Follow the instructions given.
2. Go to `https://vg.no`. Click on some articles. Go to `https://cnn.com`. Click on some articles.
3. Open the plugin
4. Explain what you see (the Dashboard)
5. Blacklist the domain vg.no
6. Explore your result
7. Explain what you see (Map)
8. See details about a country where you have polluted (Map)
9. Explore the results of the websites you have visited.
10. Find the website that have used the most energy
11. Find out how much you have polluted compared to others
12. Sign out

5 Website and Content Management System

In this appendix, present the website and the Content Management System (CMS) will be presented. The website offers information about the project and lets the user sign up for the project. The CMS makes it possible to update the information on the website without having to change the actual code. The CMS also offers a friendly interface for writing text.

Code base The code is available here: `https://github.com/Jesperpaulsen/data-collector-landing-page`

Hosting Both the Website and the CMS are hosted on Vercel[1]. Vercel offers a free hosting service for statically generated pages and server-side rendered websites. Vercel is the creator of Next.js, and the platform integrates well with Next.js.

Continuous Deployment Vercel deployment was used for Continuous deployment of both the CMS and the website. As the website is statically built, a new build of the website is required when text changes have been conducted in the CMS. On code push to the main branch, a webhook starts deploying the site and studio through Vercel. The deployment of the website can also be manually started from within the CMS.

5.1 Website

The Website is written using Next.js[2]. Next.js is a react framework for creating production websites, supporting features such as statically generated pages and server-side rendering [140]. Tailwind[3] is used for styling. When a user signs up for the project through the website, the user details are sent to the DataAnalyzer API, which in turn creates a document for the user in the database. The landing page of the website is shown in Figure 1.

5.2 CMS

For CMS, Sanity[4] is used. Sanity is a headless CMS that can be self-hosted created by a Norwegian company. The CMS is set up with different schemas defining the structure and content of the different pages. An example of one of the pages from the CMS is shown in Figure 2.

[1] `https://vercel.com/`
[2] `https://nextjs.org/`
[3] `https://tailwindcss.com/`
[4] `https://sanity.io`

Figure 1: The project's website

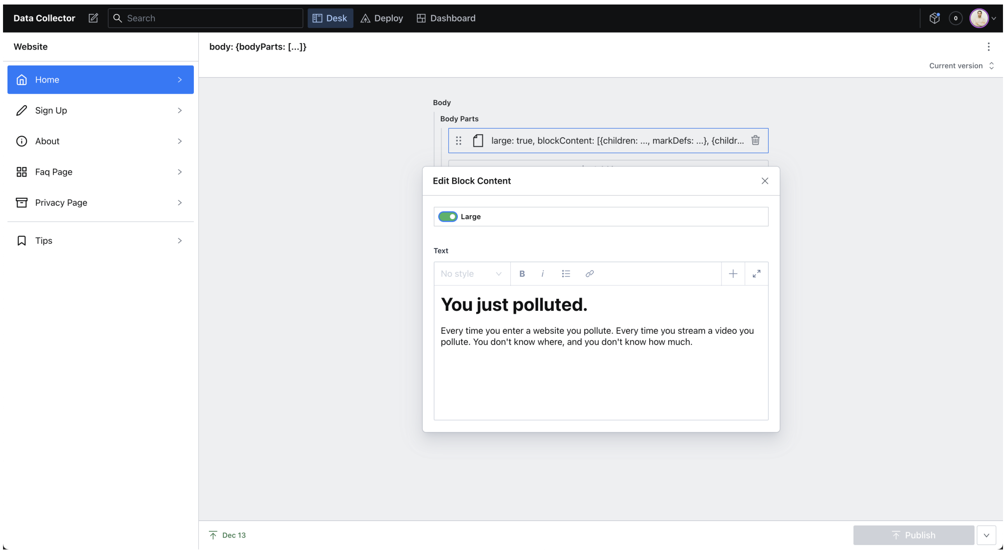

Figure 2: The project's CMS

6 Tips and challenges presented to the user during the project

During the project, the participants received tips on how to reduce their digital climate footprint through the DataPresenter. They also received a challenge to reduce their digital climate footprint. The tips and challenges are presented in Table 1.

ID	Tips
1	The more data you use, the more you pollute. The internet is physical. Every byte of data you use travels in cables from and to data centers around the world. Most of the data centers are located in the USA, and the data to/from Europe travel under the Atlantic ocean. To reduce your climate footprint, reduce the amount of data you use. Lower the quality of videos you watch. Don't watch videos you don't need to watch. Download music, instead of streaming it over and over again. A map of the submarine internet cables can be found here: https://www.submarinecablemap.com/
2	Everything that's free on the Internet is using the world's resources. Advertisements often finance the free services. Advertisements in the form of videos are polluting. Advertisements in the form of images are also polluting. There is no such thing as a free lunch when using the Internet. The earth is paying for the lunch. Stop watching videos with advertisements unless you really need to watch the video.
3	The cloud isn't a cloud. The cloud is datacenters all around the globe. Unlike your hard disk, the cloud doesn't know when you want to access your files. Therefore, they are always available. Constantly using energy. If you want to reduce your digital climate footprint, you should (1) delete the files you don't need anymore and (2) move the files you don't need to access too often to a hard drive.
4	The more people you follow on a social media platform, the more you will pollute when using the platform. The more followers you have when you post an image or a video, the more pollution you will cause. Stop following accounts you don't care about. Stop posting images, videos, and stories too often. Don't post an image or video unless you think it will value you or your followers. Liking or commenting on a post that's posted by an account with many followers will cause more pollution. Consider if you really need to give the post a like or a comment.
5	Research shows that between 80% to 90% of all created data is never used after its creation. It's created, stored, and never used again. If it's stored in the cloud, it is constantly using energy. Do yourself a favor, stop producing waste. In comparison to plastic, digital waste is relatively easy to recycle. Do the earth a favor, reduce the digital climate footprint of the data you create.

6	Now that you have learned that 80-90% of all created data never is used after its creation, think about your own data. Do you have any data stored in the cloud you know you don't need to have stored? Make a positive climate impact, delete at least 20 files, images, or videos from the cloud.
7	Research shows that only 20% of the pollution that is caused by a device connected to the Internet comes from its lifetime. The other 80% comes from the production of the device. The Internet components are getting more effective, processing more and more data, using less power per byte. But, at the same time, the production of the devices is consuming more and more resources and energy. The Internet components only have an expected lifetime of three to five years. We need to ensure that the digital components can live longer. We need to reduce the amount of data that we use so that older Internet components extend their lifespan. Only by doing so will we be able to reduce the Internet's climate footprint.
8	Delete apps you don't use on your phone. Whenever an app is automatically updated through App Store or Google play, the whole app is first deleted before the new version of the app is downloaded. This is done on every mobile that has the app downloaded. A single update of a popular app will result in a lot of data being sent across the Internet and, therefore, more pollution.
9	The latest research shows that you are able to reduce your carbon footprint by 86% if you stream videos in standard definition instead of high definition. By turning off your camera in a video meeting you are able to reduce your carbon footprint by 96%. Next time you are in a digital meeting, consider if it's necessary to have the camera turned on.

Table 1: Tips and challenges presented to the participants

7 Questionnaires

In this section the two questionnaires will be presented.

7.1 Questionnaire 1

Part 1: Basic information The first part of the questionnaire collected background information about the participant.

ID	Question	Type
1.1	What is your e-mail address?	Open question
1.2	What is your age?	Radio buttons
1.3	What is your gender?	Dropdown
1.4	Which country are you from?	Dropdown
1.5	Have you developed or designed websites and/or apps that use the internet?	Yes/No
1.6	Have you studied, or are you studying, how to develop or design websites and/or apps that use the internet?	Yes/No

Part 2: Information about your Internet usage The second part examined how the participants used the Internet.

ID	Question	Type
2.1	On a daily basis, how many hours do you spend streaming music on platforms like Spotify and Soundcloud?	Dropdown
2.2	On a daily basis, how many hours do you spend on platforms like TikTok and YouTube?	Dropdown
2.3	On a daily basis, how many hours do you spend on platforms like Facebook, LinkedIn, and Twitter?	Dropdown
2.4	On a daily basis, how many hours do you spend on streaming platforms like Netflix and HBO Max?	Dropdown
2.5	On a daily basis, how many hours do you spend in digital meetings on platforms like Zoom and Teams?	Dropdown
2.6	On a daily basis, how many hours do you spend on regular browsing?	Dropdown
2.7	On a daily basis, how many hours do you spend on online gaming?	Dropdown

Part 3: Background questions The third part, two background questions about the participant's knowledge about the energy usage of the Internet.

ID	**Question**	**Type**
3.1	Before entering this study - did you know that the Internet pollutes?	Radio buttons
3.2	Compared to the aviation industry, how much pollution do you think that Internet usage causes?	Radio buttons

Part 4: The Internet and the Environment The fourth part of the first questionaire focused on the environment and the Internet.

ID	**Question**	**Type**
4.1	I care about the environment	Likert scale
4.2	I know how the "cloud" works	Likert scale
4.3	It is important for me to make environmental friendly choices	Likert scale
4.4	I want to learn how the Internet pollutes	Likert scale
4.5	I know how I can reduce my digital carbon footprint	Likert scale
4.6	I want to learn how I can reduce my digital carbon footprint.	Likert scale

Part 5: The Environment The fifth part examined how much the participant think about the environment in general, and when using the Internet.

ID	Question	Type
5.1	I feel guilty about flying.	Likert scale
5.2	I feel guilty about using the Internet.	Likert scale
5.3	I think about the environment before I visit a website	Likert scale
5.4	I think about the environment before I post an image to social media	Likert scale
5.5	I think about the environment before I watch a video on the Internet	Likert scale
5.6	I think about the environment before I stream a video	Likert scale
5.7	I think about the environment before I upload a file to the cloud	Likert scale

Part 6: Willingness to change The sixth part survey the participants will to change their behavior.

ID	Question	Type
6.1	I'm willing to change my behavior on the Internet to save the environment	Likert scale
6.2	I'm willing to reduce the amount of data I'm using on the Internet if it will reduce my digital carbon footprint	Likert scale
6.3	I'm willing to reduce the time I spend on the Internet if it reduces my digital carbon footprint.	Likert scale
6.4	I'm willing to have less stored in the cloud if it will reduce my digital carbon footprint	Likert scale
6.5	I want to learn where in the world I pollute when using the Internet	Likert scale
6.6	I'm willing to stop visiting websites that pollute much if I get feedback on how much they pollute	Likert scale
6.7	I feel images, videos and animations are important to make websites more interesting	Likert scale
6.8	I'm willing to visit webpages that looks less "fancy" and uses less images and videos if it will reduce my digital carbon footprint	Likert scale
6.9	I wish search engines like Google sorted the relevant search results based on how much they pollute	Likert scale
6.10	I wish cloud providers made it easier to delete images, videos, and documents.	Likert scale

Part 7: Questions for developers and designers The seventh part consisted of questions specific for developers and designers. Question 7.1 and 7.2 was visible if the participant answered *Yes* on question 1.6. Question 7.3 was visible if the participant answered *Yes* on question 1.5.

ID	Question	Type
7.1	While studying, I learned about how the internet pollutes.	Likert scale
7.2	While studying, I learned how I could create websites and apps that pollute less.	Likert scale
7.3	When creating or designing websites and/or apps, I think about how to create websites and/or apps that pollute less.	Likert scale

7.2 Questionaire 2

Part 1: Basic information The first part of this questionaire only collected the email of the participant.

ID	Question	Type
1.1	What is your e-mail address?	Free text

Part 2: Information about your Internet usage The second part examined if the participants felt like their internet behaviour had changed after being part of the project.

ID	Question	Type
2.1	After joining the project, I spend less time streaming music on platforms like Spotify and Soundcloud	Likert scale
2.2	After joining the project, I spend less time on platforms like TikTok and YouTube	Likert scale
2.3	After joining the project, I spend less time on platforms like Facebook, LinkedIn, and Twitter	Likert scale
2.4	After joining the project, I spend less time on streaming platforms like Netflix and HBO Max	Likert scale
2.5	After joining the project, I turn off my camera more often in digital meetings on platforms like Zoom and Teams	Likert scale
2.6	After joining the project, I spend less time on regular browsing	Likert scale
2.7	After joining the project, I spend less time on online gaming	Likert scale
2.8	After joining the project, I think about how I can reduce my digital carbon footprint when using the Internet	Likert scale

Part 3: Plugin Features The third part examined the different features of the plugin and their perceived value. The description of this part was: *"For the next questions, "Useful" means that it helped you become more aware of your digital climate footprint. In this context, pollution means CO2 equivalents. A CO2 equivalent is defined as: "A carbon dioxide equivalent or CO2 equivalent is a metric measure used to compare the emissions from various greenhouse gasses on the basis of their global-warming potential (GWP), by converting amounts of other gases to the equivalent amount of carbon dioxide with the same global warming potential.""*

ID	**Question**	**Type**
3.1	It was useful to recieve tips on how I can reduce my digital carbon footprint	Likert scale
3.2	It was useful to see my usage and pollution caused in real-time	Likert scale
3.3	It was useful to see my daily pollution caused compared to my average pollution, as well as others average pollution (In percentage)	Likert scale
3.4	It was useful to compare my energy consumption to the distance I could have been able to drive an electric car with the same energy consumption	Likert scale
3.5	It was useful to compare my pollution to the distance I could have been able to drive a petrol car, to generate the same amount of pollution	Likert scale
3.6	It was useful to get push notifications with pollution hints and updates	Likert scale
3.7	It was useful to get reports on my pollution generated from Internet activity previous days	Likert scale
3.8	It was useful to be able to explore a map to see where in the world I pollute	Likert scale
3.9	It was useful to have a chart with the average usage of everyone else compared to my usage	Likert scale

Part 4: The Internet and the Environment The fourth part of the second questionaire had two of the questions from the fourth part of the first questionaire.

ID	**Question**	**Type**
4.1	I care about the environment	Likert scale
4.2	I know how the "cloud" works	Likert scale
4.3	I know how I can reduce my digital carbon footprint	Likert scale

Part 5: The Environment The fifth part in the second questionaire was identical to the fifth part of the first questionaire.

ID	**Question**	**Type**
5.1	I feel guilty about using the Internet.	Likert scale
5.2	I think about the environment before I visit a website	Likert scale
5.3	I think about the environment before I post an image to social media	Likert scale
5.4	I think about the environment before I watch a video on the Internet	Likert scale
5.5	I think about the environment before I stream a video	Likert scale
5.6	I think about the environment before I upload a file to the cloud	Likert scale

Part 6: Final questions The sixth part was a summary of the user's experience of the project.

ID	Question	Type
6.1	Based on the feedback you got through the plugin, what do you experience with your own digital carbon footprint?	Radio buttons
6.2	The plugin helped me reduce my digital carbon footprint	Likert scale
6.3	The plugin helped me become more aware of I pollute when I use the Internet	Likert scale
6.4	I have a better understanding of how the Internet works	Likert scale
6.5	I have changed my Internet habits	Likert scale
6.6	Did you do any of the following to reduce your digital carbon footprint?	Checkboxes
6.7	Number of files you have deleted from the cloud after joining the project	Radio buttons
6.8	Do you have any other feedback on either the plugin or the project you wish to give?	Free text

8 Results from the Experiment

8.1 Line diagrams

In this section, the data collected from each user is presented. Each user have the same color for each of the diagrams.

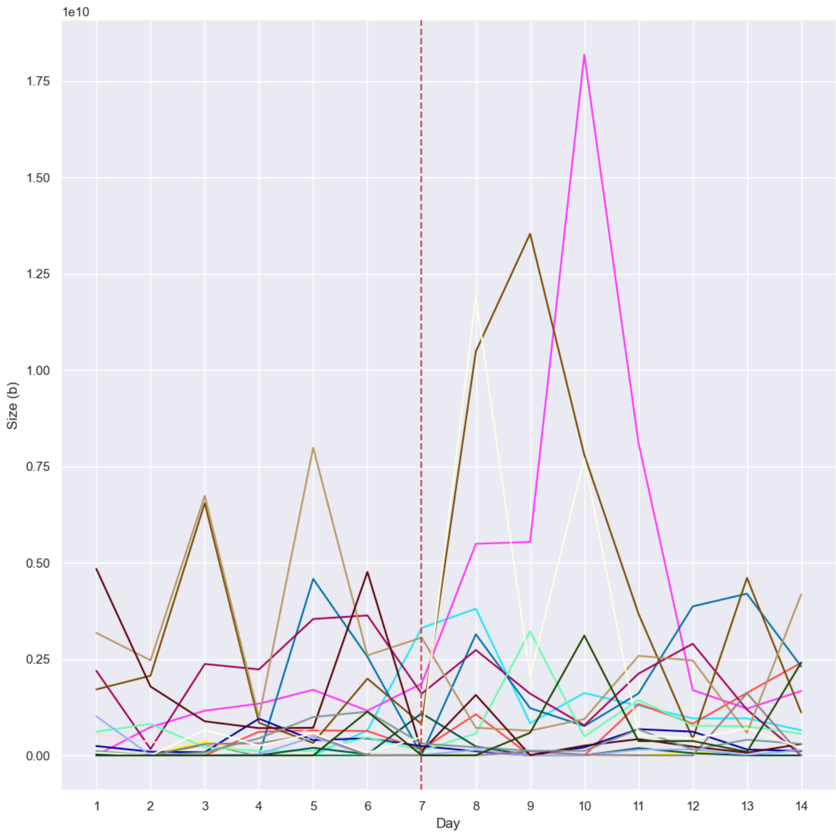

Figure 3: Size line diagram for each user. The red dotted line indicates when the users got access to see their usage.

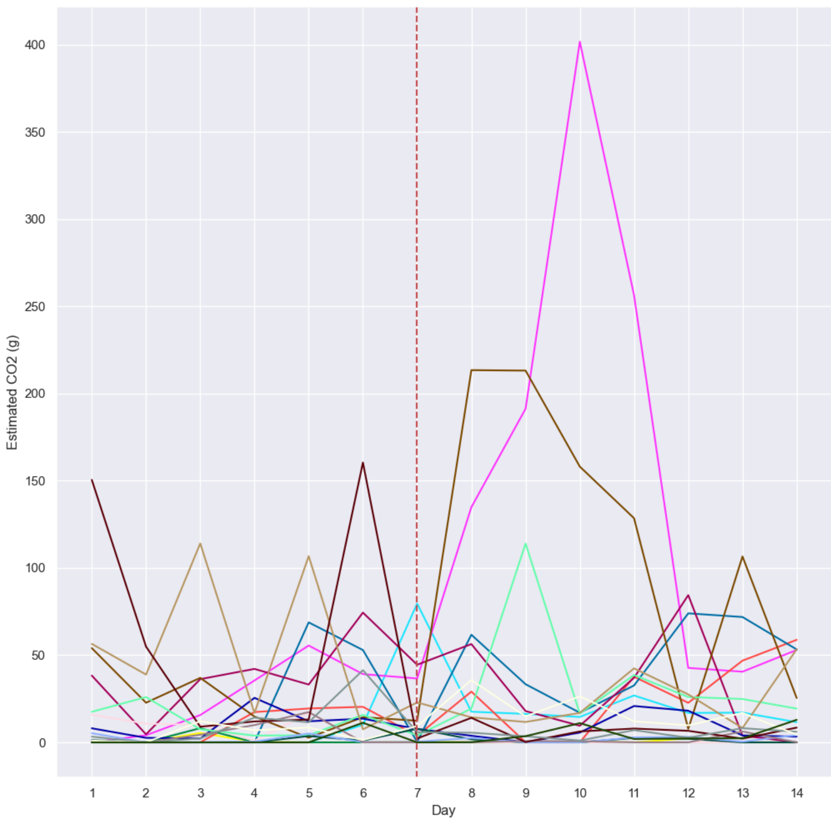

Figure 4: CO_2 line diagram for each user. The red dotted line indicates when the users got access to see their usage.

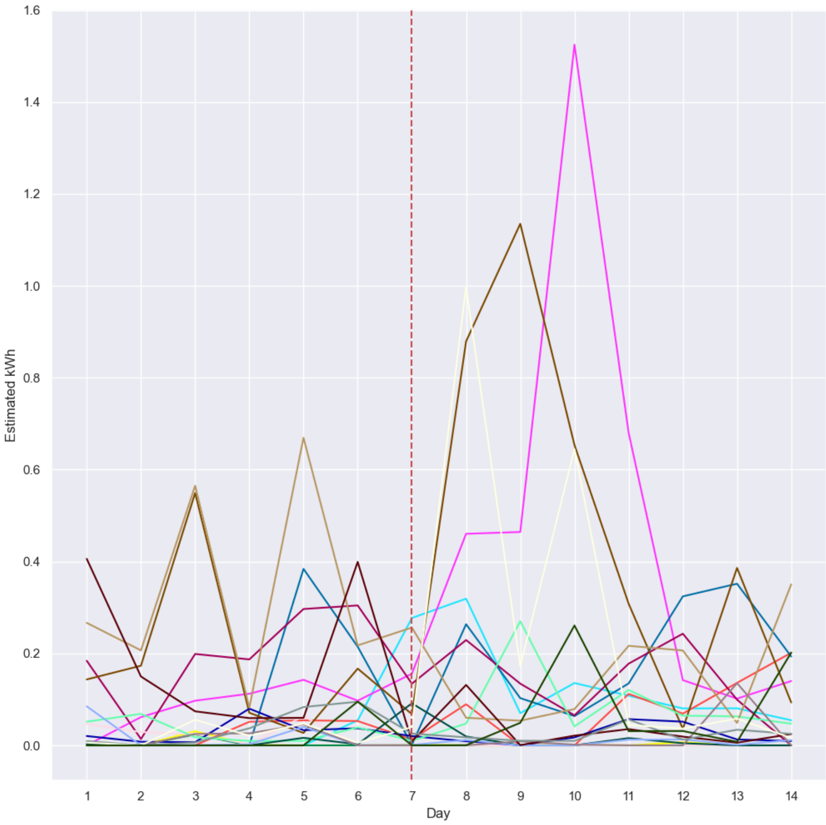

Figure 5: kWh line diagram for each user. The red dotted line indicates when the users got access to see their usage.

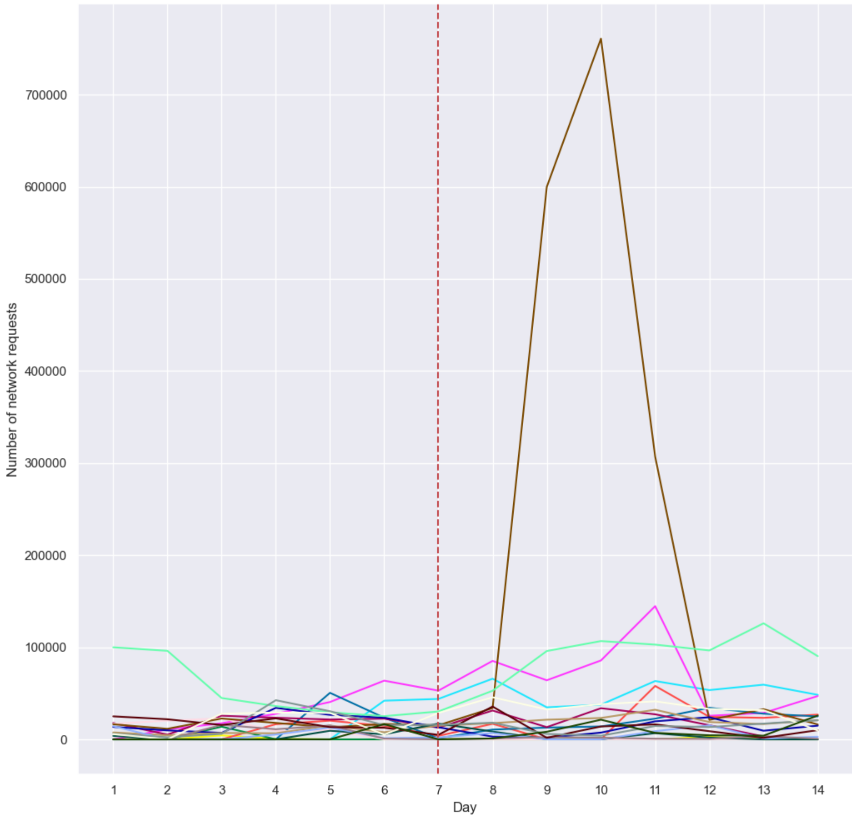

Figure 6: Number of calls line diagram for each user. The red dotted line indicates when the users got access to see their usage.

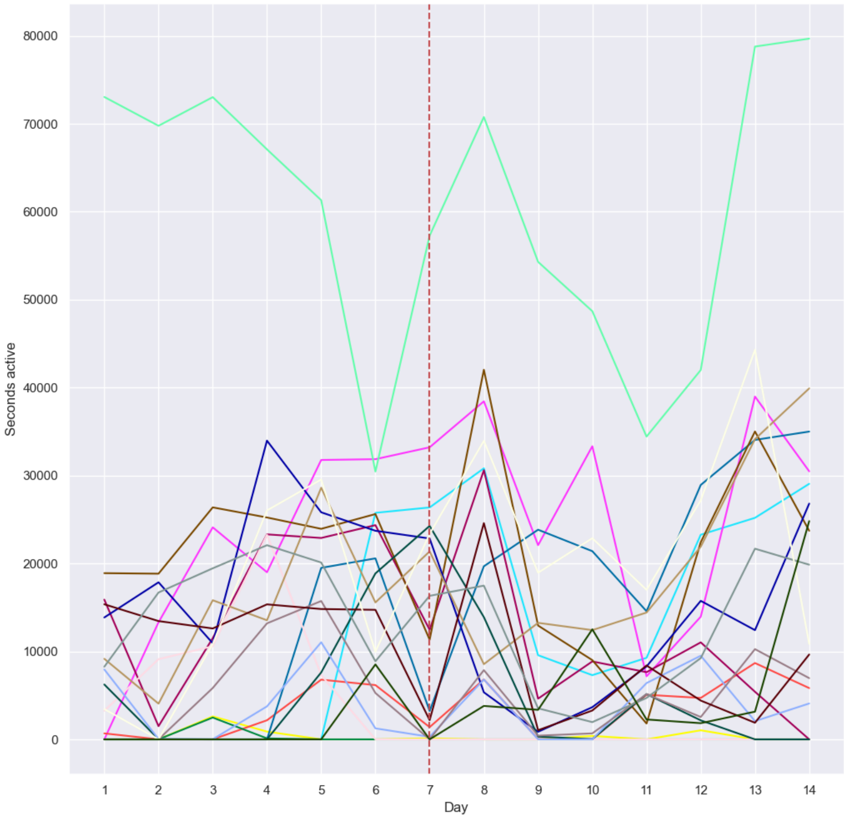

Figure 7: Seconds active line diagram for each user. The red dotted line indicates when the users got access to see their usage.

8.2 Scatter diagrams

In this section, scatter diagram for the different attributes per day is presented to give a clearer overview of the usage per day. The blue line is the average per day.

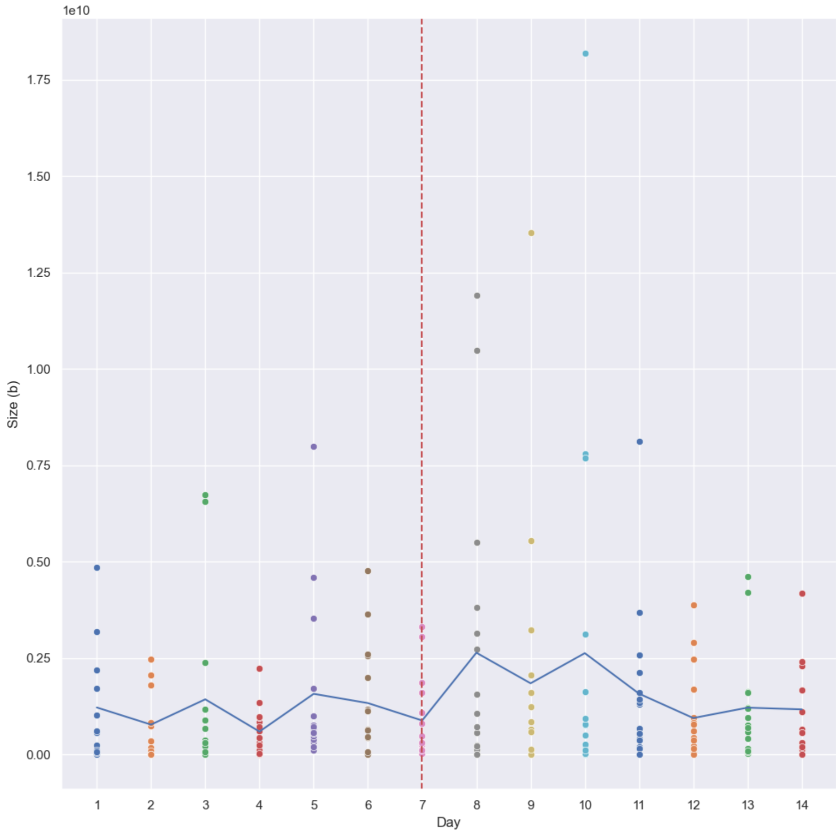

Figure 8: Size scatter diagram. The red dotted line indicates when the users got access to see their usage. The blue line is the average per day.

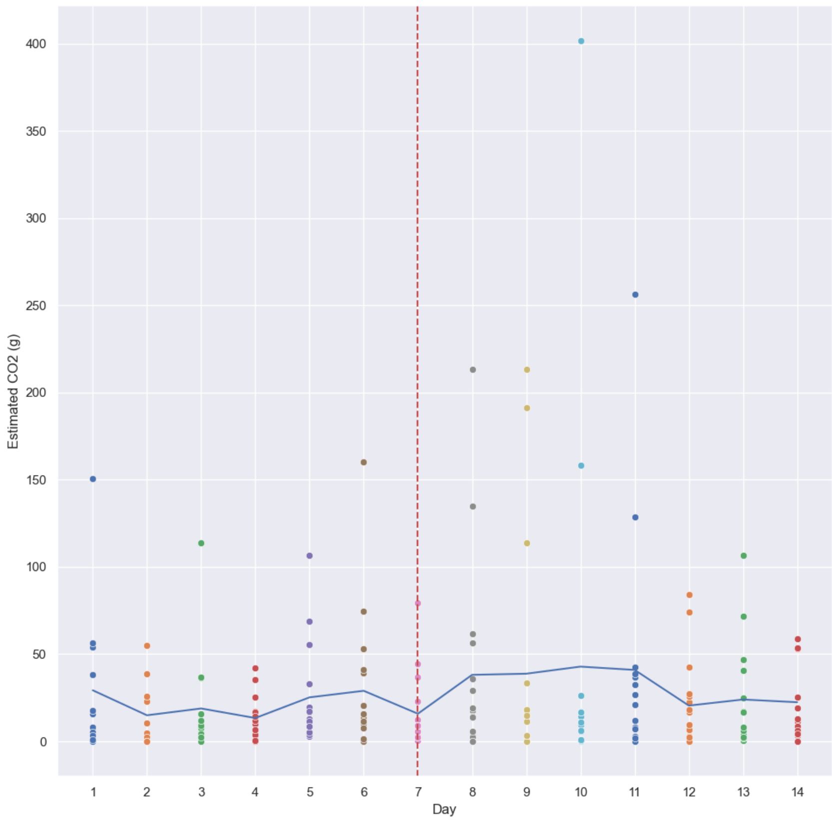

Figure 9: CO_2 scatter diagram. The red dotted line indicates when the users got access to see their usage. The blue line is the average per day.

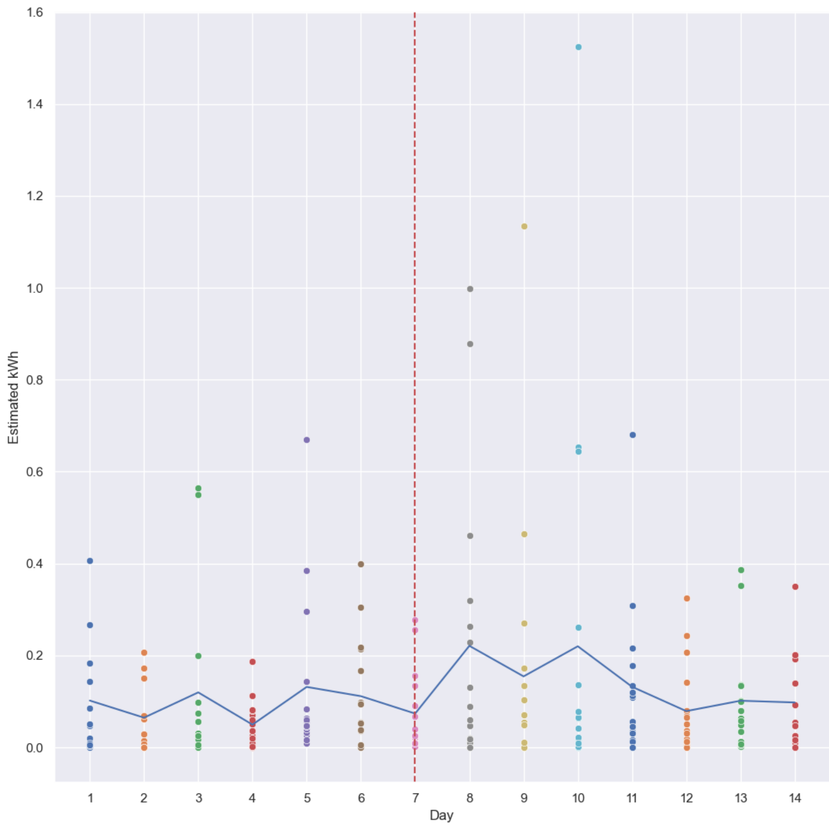

Figure 10: kWh scatter diagram. The red dotted line indicates when the users got access to see their usage. The blue line is the average per day.

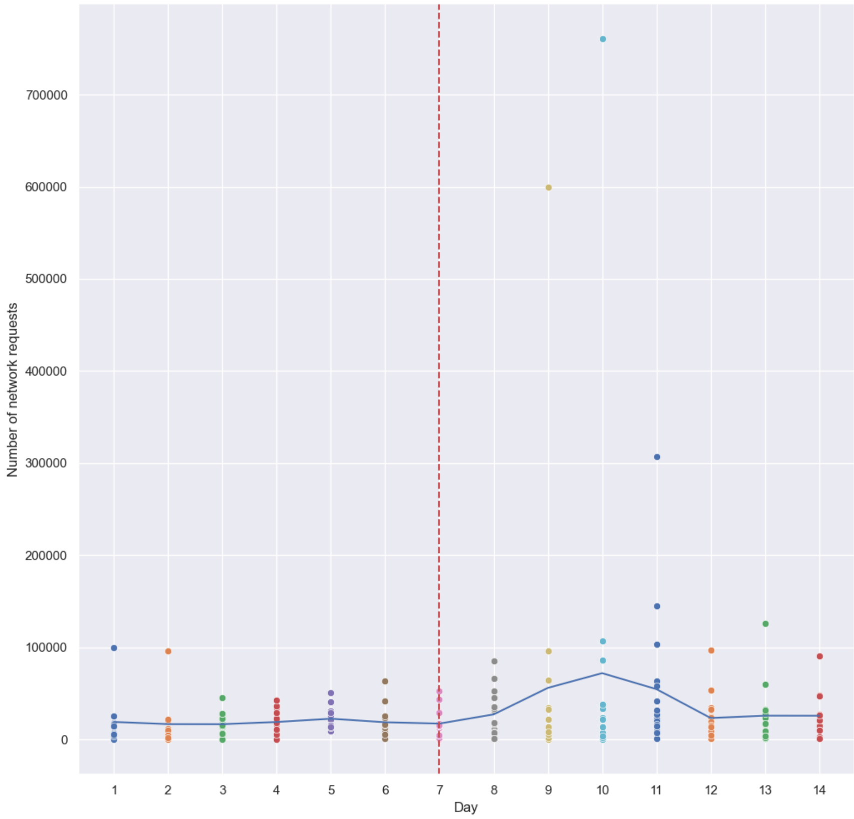

Figure 11: Number of calls scatter diagram. The red dotted line indicates when the users got access to see their usage. The blue line is the average per day.

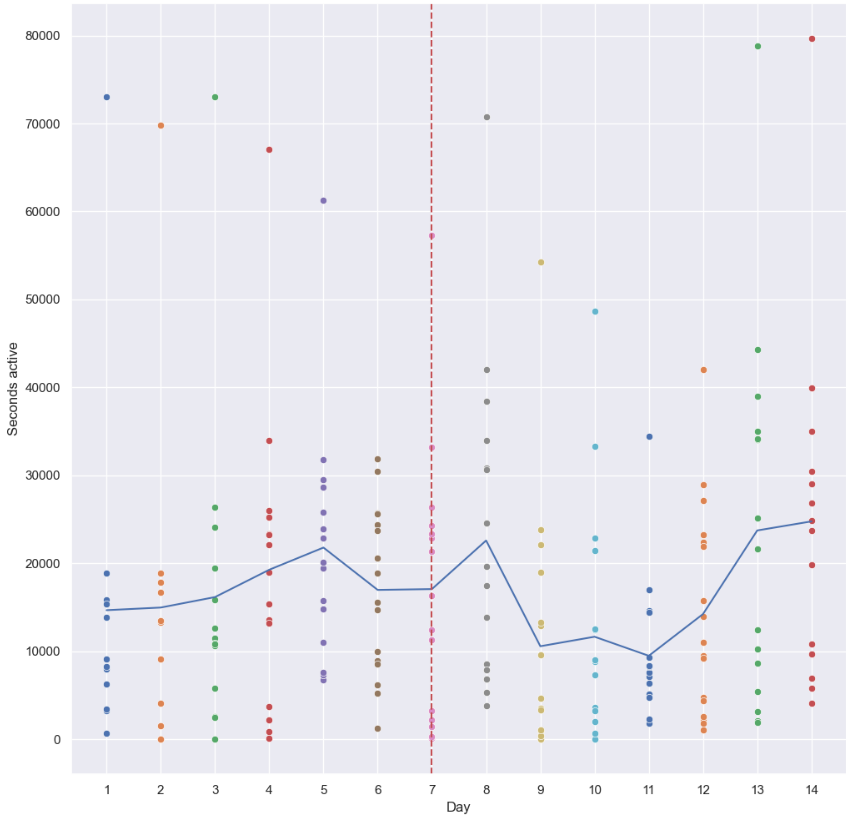

Figure 12: Seconds active scatter diagram. The red dotted line indicates when the users got access to see their usage. The blue line is the average per day.

8.3 Box diagrams

In this section, box diagram for the different attributes per day is presented.

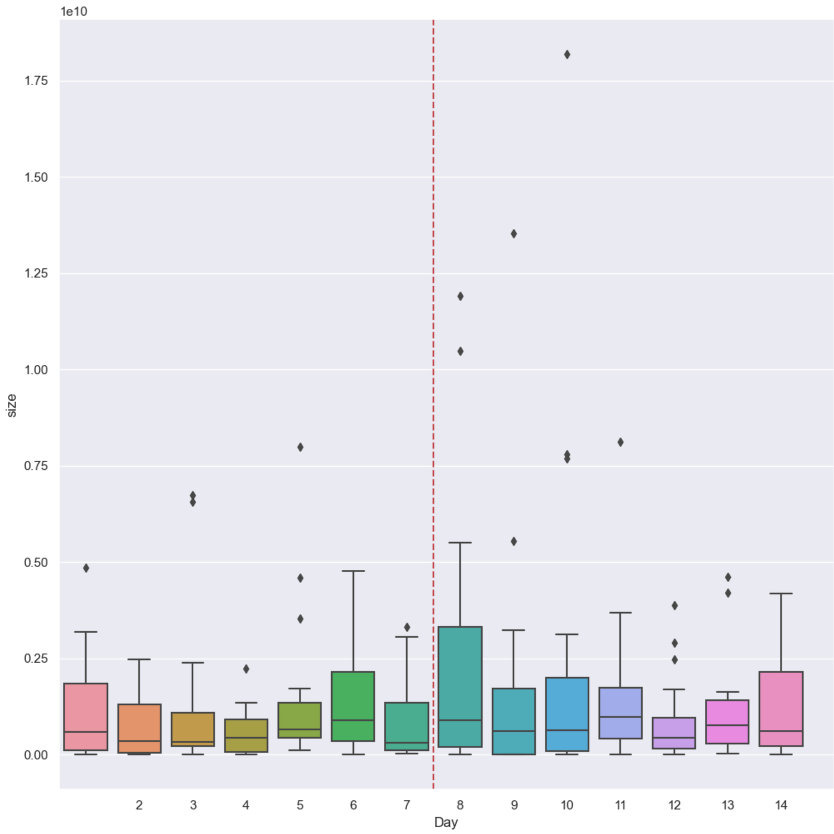

Figure 13: Size box diagram. The red dotted line indicates when the users got access to see their usage.

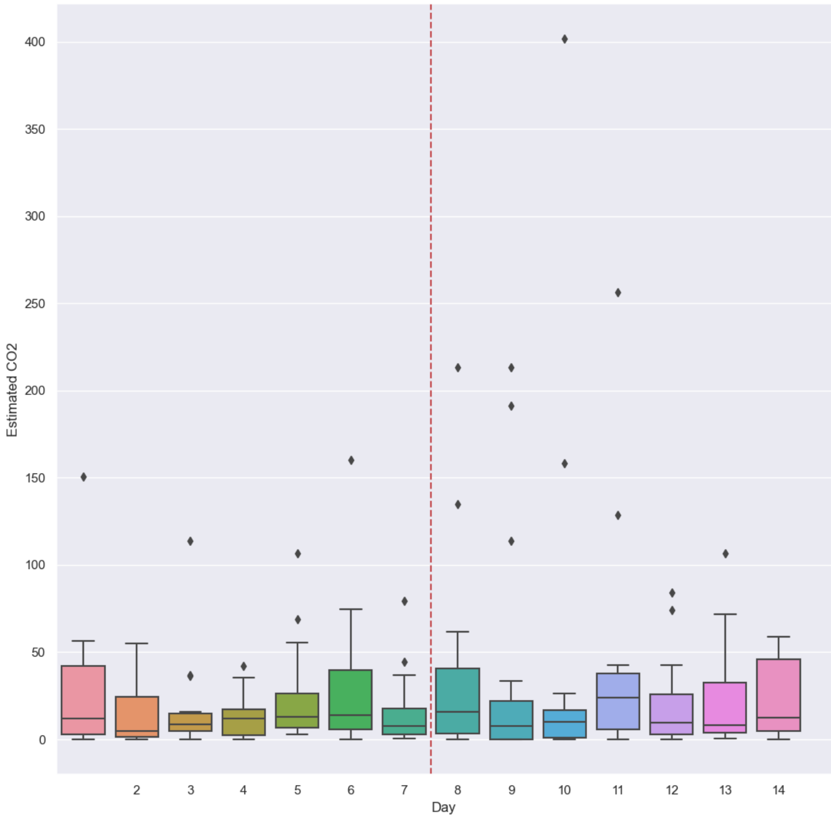

Figure 14: CO_2 box diagram. The red dotted line indicates when the users got access to see their usage.

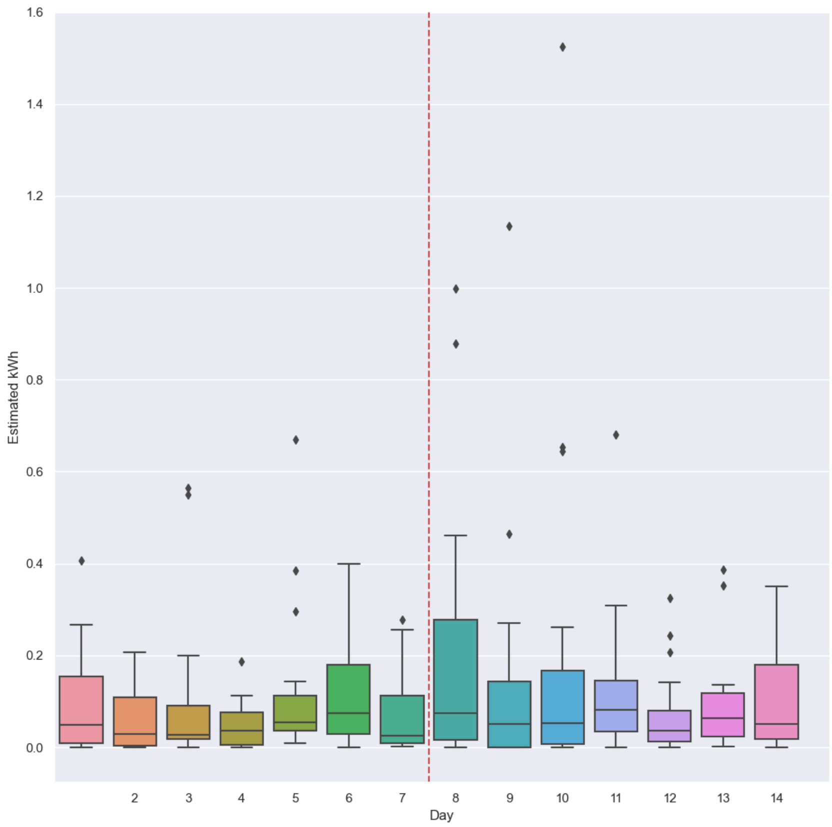

Figure 15: kWh box diagram. The red dotted line indicates when the users got access to see their usage.

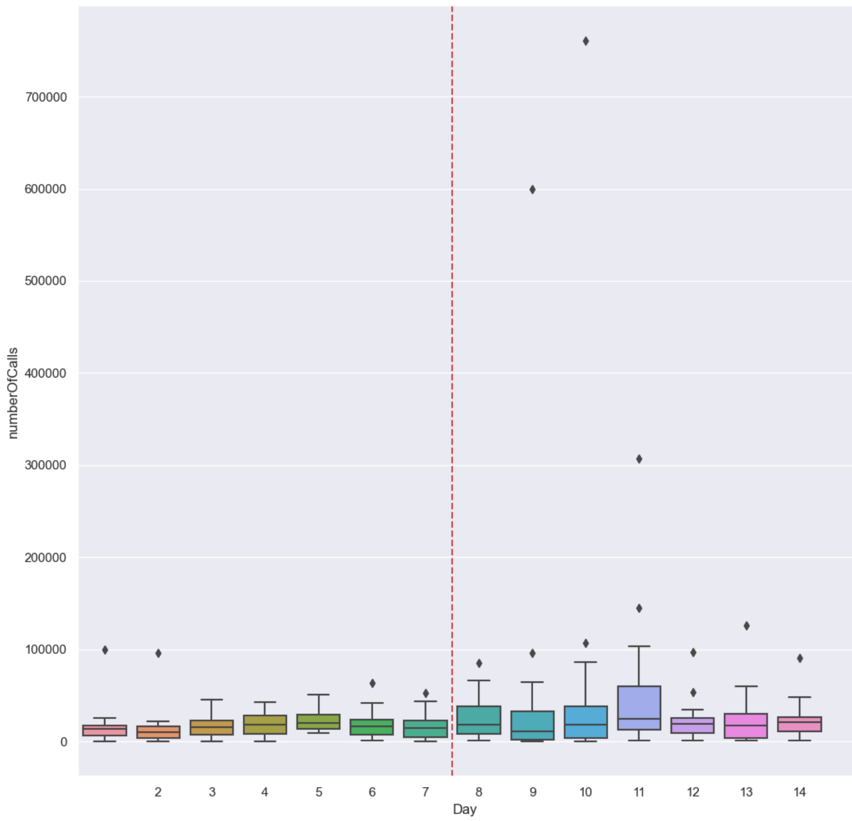

Figure 16: Number of calls box diagram. The red dotted line indicates when the users got access to see their usage.

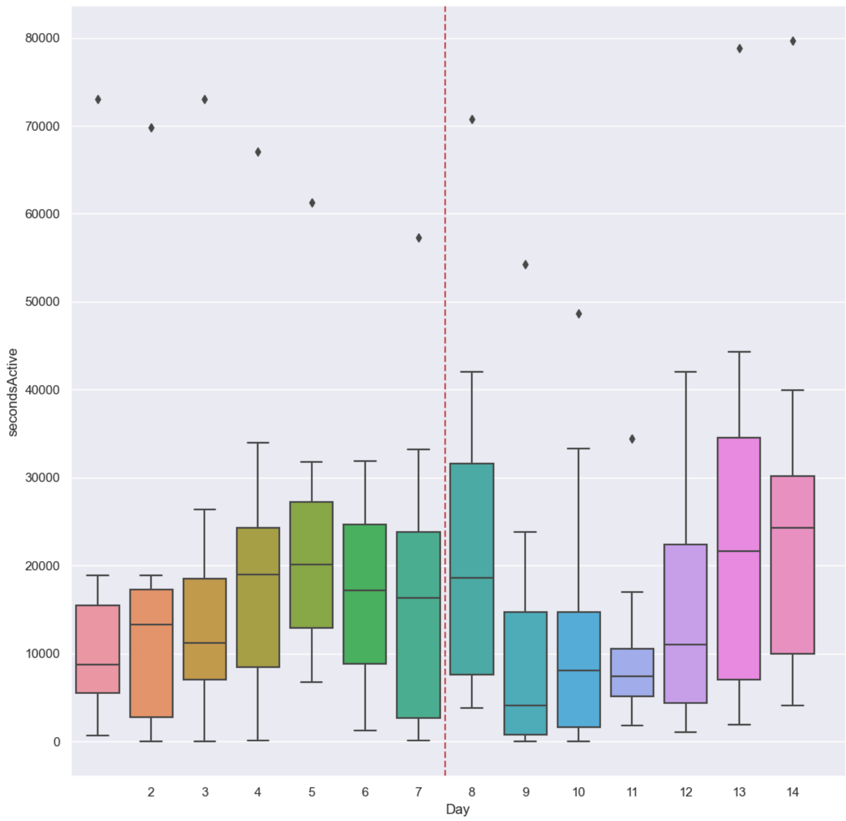

Figure 17: Seconds active box diagram. The red dotted line indicates when the users got access to see their usage.

Jesper G. Paulsen

Making the Invisible Visible - How people can become aware of their own digital climate footprint to save the earth

NTNU
Norwegian University of
Science and Technology