Summary

This project explores how a physical representation of cookie data can encourage reflection on data privacy in a digital environment. Cookies collect large amounts of data, often in ways that can be hard for users to understand. Additionally, many websites use deceptive design practices to encourage data collection, which often leaves users feeling overwhelmed and powerless. This is known as privacy fatigue, which can lead users to disclose personal information despite privacy concerns. In addition to privacy fatigue, there tends to be a contradiction between how users view data privacy and their interactions with cookies banners. This behavior reflects the privacy paradox. Research into data physicalization and tangible designs revealed how it is possible to make data perceptible through a tangible form, enabling users to gain a deeper understanding of data. Existing studies show how making abstract digital processes more visible and interactive, can help support self-reflection and increase user awareness. This exploration led to the development of DataCrumb: a physical prototype that visualizes data being shared through lights, sounds, and a display. With DataCrumb, the aim is to make data collection more tangible and encourage users to reflect on their digital privacy, addressing the abstract and often unclear nature of cookies and data revealed in the initial interviews.

To evaluate whether the prototype fulfilled its intended purpose, three test sessions were conducted. The tests revealed that the participants became more aware of how data was being collected. The physical nature of DataCrumb made it hard to ignore, which led to reflections on the participants' data privacy.

Overall the project suggests there is potential in making cookie data physical, as it can foster a deeper understanding and encourage users to engage more critically with their digital privacy.

Project Title:
DataCrumb: A Tangible Approach to Visualizing Cookie Data

Field of Study:
Digitalisation and Application Development (DAD)

Educational Institute:
Aalborg University (AAU)

Project Group:
Group 1

Project Period:
Master Thesis (10th Semester)

Group Members:
Aisho Mohamed Ali (20183380)
Anna Stentoft Ries (20205031)
Katrina Hvítklett (20221589)

Supervisor:
Florian Echtler

Number of Pages:
14

Date of Completion:
10.06.2025

DataCrumb: A Tangible Approach to Visualizing Cookie Data

Anna Stentoft Ries
Aalborg University
aries20@student.aau.dk

Aisho Mohamed Ali
Aalborg University
amoha18@student.aau.dk

Katrina Hvítklett
Aalborg University
khvitk22@student.aau.dk

Figure 1. Interaction with DataCrumb

Figure 2. Top view of DataCrumb

Figure 3. Side view of DataCrumb

Abstract

Most people simply accept cookies, leaving behind a trail of personal data. However, the overwhelming amount of information on cookie banners and policies makes it difficult for individuals to fully understand and manage their data, often leading to feelings of helplessness and disengagement. To address this, we created DataCrumb: a tangible prototype that physicalizes data sharing through lights, sounds, and a display. Our initial research identified issues of the privacy paradox and privacy fatigue, which complicate users' ability to navigate data privacy effectively. DataCrumb was designed to increase awareness and encourage reflection on data sharing that occurs during everyday online activities. We evaluated the prototype through a three-part test: a pre-interview, an in-home testing session, and a post-interview. The results showed that DataCrumb raised participants' awareness of data and cookie usage and prompted greater reflection on their online behaviors.

1 Introduction

Data collection has become more prevalent in today's digital landscape, and it is occurring every time users enter a website or application, often through the acceptance of cookies. Cookies, which are small text files that store data on a user's device, function by assigning a unique identifier that allows websites to recognize returning users [26]. The International Chamber of Commerce UK defines four different types of cookies: *necessary*, *performance*, *functional* and *tracking*, each serving a different purpose [19]. Necessary cookies ensure that a website operates correctly, while performance cookies collect anonymous usage data, and functional cookies store user preferences to enable personalized features [19]. Tracking cookies collect information about a user's browsing habits and can be used in targeted advertising and cross-site tracking [19]. In addition to the four types of cookies, cookies can be divided into fist-party and third-party cookies. First-party cookies are set by the website being visited and are typically used for functionality and user experience enhancement. In contrast, third-party cookies are set by external parties that the user may not directly interact with. These cookies track users' online behavior across multiple sites, often without the user knowing which companies have access to their information [21]. As a result, cookies have become a valuable source of personal data. This makes them subject to the General Data Protection Regulation (GDPR) [6, 21]. GDPR specifies that users need to give explicit consent before websites can use cookies, with the exception of 'strictly necessary' cookies [17, 26]. The EU ePrivacy Directive, often referred to as the 'Cookie Law,' supports GDPR by setting specific rules on cookie consent and requiring websites to inform users about data collection [17]. This regulation has been implemented to ensure data transparency and to control cookie usage, which protects user data and ensures that companies handle information responsibly [6].

In this project, we explore how the collection of personal data through cookies impacts users and how the materialization of data can facilitate self-reflection. Drawing inspiration from studies on data physicalization and tangible design, we developed DataCrumb: a prototype that aims to make the abstract concept of data more tangible, encouraging users to reflect on their data privacy. The project begins with an exploration of related work, including deceptive design practices and the effects of privacy fatigue and the privacy paradox.

In this section, we also explore how data physicalization and tangible design approaches have been used to support user awareness and understanding of data. In addition to related work we present our research question, and in a method section, our interview findings, our ideation process and the final solution. The Implementation section explains the technical setups and pseudocode to outline how the prototype has been implemented. In the project we will also present the findings from our test sessions, and how these tests were conducted. Finally, we discuss the project's limitations and the potential for future development.

2 Related work

2.1 Deceptive Design: Influencing User Decisions

Privacy regulations related to data protection, such as the EU ePrivacy Directive and GDPR, have caused an increase in the use of cookie banners on websites [26]. Ideally, cookie banners should provide a clear choice for the user on whether to accept or reject cookies, expressing their consent on which data is collected [18]. While GDPR has led to an increase in data transparency and control for users, many websites often use deceptive design to encourage data collection [11]. Since it is in the companies' interest to collect user data, they often use text-heavy banners and implement nudges, which guide users toward actions that primarily benefit the company rather than the user [3, 20, 21]. These nudges are also referred to as *dark patterns*, as they undermine principles of the EU ePrivacy Directive and GDPR [11].

Dark patterns manipulate users into accepting being tracked while creating an illusion of control [11]. Research shows that design influences consent more than trust, with an overwhelming amount of text and elements like button placement, forced choices, and colors heavily influencing user decisions [21]. As a result, cookie banners often mislead rather than inform, undermining informed consent and prioritizing data collection over user privacy. This raises ethical concerns about privacy and consent, as it undermines the user's ability to make informed decisions about their online privacy.

2.2 Privacy Fatigue & Privacy Paradox

Even though multiple studies emphasize the influence dark patterns can have on user's interactions with cookie banners, research shows that users feel little control over their personal data, regardless of banner design [11]. *Privacy fatigue* could be a contributing factor to the feeling of loss of control [7, 31]. While privacy concerns make users more cautious of their online behavior and sharing of data, users experiencing privacy fatigue tend to feel resigned and disengaged from the issue [31]. Continuous consent requests and complex privacy policies can make users feel overwhelmed and more likely to disengage from privacy decisions. As a result, users experiencing privacy fatigue tend to disclose personal information online, even if they have privacy concerns [7]. This can explain why users make decisions regarding their privacy that contradicts their concerns and intentions, which supports the *privacy paradox*.

The privacy paradox is the phenomenon where users claim to value privacy, but still disclose personal data and fail to use measures to protect their privacy. This inconsistency between intentions and behavior does not mean users do not care about privacy. Rather, privacy decisions are made in specific situations, while privacy concerns reflect broader, more general attitudes [30]. In relation to cookie data, the privacy paradox is demonstrated when users accept all cookies, even though they have expressed concerns with online privacy. A study by Mejtoft et al. (2023) found that there is a mismatch between the participants' disclosed behavior and their observed behavior, when presented with cookie banners, as they tend to accept more cookies than they say they do [21]. Additionally, the paper "(Un)informed Consent: Studying GDPR Consent Notices in the Field" indicates that users are frustrated with cookie policies or unsure about their impact [33]. Therefore, users often ignore cookie banners, even though they state that their interactions are privacy focused.

2.3 Data Physicalization

Data physicalization, as defined by Dragicevic et al. in "Data Physicalization" is a "*physical artifact whose geometry or material properties encode data*" [8]. The goal of data physicalization is to create "*interactive, visual representations of abstract data to amplify cognition*" [8]. These representations can serve a wide range of purposes in different contexts, like analytics, communication, education or self-reflection. By making data perceptible through a tangible form, data physicalization can help people explore and gain a deeper understanding of data [15].

In the context of cookies and online privacy, data physicalization also holds potential for promoting self-reflection and awareness [8]. As data collection has become increasingly embedded in digital interactions, users are often left with little transparency and control over how their personal information is tracked, stored and used [7, 31]. So, while many claim to value their privacy, their behavior can contradict those concerns [30]. With these challenges in mind, data physicalization can help represent data and encourage self-reflection. Such introspection can, for example, confront users with their behavior, which in turn can lead to a clearer and more understandable view of their actions[8]. An example of this can be seen in the study "SweatAtoms" which transformed participant's activity data, including steps and heart rate, into a unique sculpture to promote self-reflection [16]. In another study, the well-being of a plant served as a representation of physical activity, emphasizing the relationship between the participant and their plant [4]. By turning data into physical artifacts, these two studies were able to prompt participants to reflect on their own behavior. Making data physical and allowing people to see, interact or touch it through an artifact, can help create a better understanding of data [8].

2.4 Tangible Designs in Research

Prior research in tangible design has investigated how physical artifacts can be used to communicate and contextualize data. In regards to data collection and online privacy, various studies have explored how tangible designs can make the abstract concept of data collection more comprehensible for users [14].

One study "Dataslip: Into the Present and Future(s) of Personal Data" highlights the issue of personal data through a provocative, interactive ATM-like installation [10]. Participants interact with the installation by answering five questions, whereafter it prints out long, detailed receipts of collected data based on their answers. The receipt is printed out slowly to foster anticipation and showcase the creepiness of information gathered. The study showed that the participants in general were concerned by the amount of information being printed, emphasizing how unaware people are of what data is collected. The study also revealed various challenges and concerns, such as the privacy trade-off, where users want to protect their data while simultaneously using platforms that collect it [10]. The paper demonstrates the lack of data transparency, where users do not fully understand what they are consenting to, when accepting cookies. By providing the participants with a tangible representation of their personal data, Dataslip effectively enforces the extent of data collection, to promote awareness and reflection [10].

Another example of a study that addresses this problem is “PrivacyCube: A Tangible Device for Improving Privacy Awareness in IoT” [22]. The PrivacyCube is a physical artifact that is designed to help users manage data privacy in Internet of Things (IoT) environments. The data is represented by icons that light up when data is being collected. The PrivacyCube being tangible and placed in the homes of the families using it, forces them to reflect on how and when their data is being collected. By enabling them to engage with a physical cube, which has been previously used as a learning tool of complex concepts, users can more easily gain an understanding of data practices and the management of their personal information [22].

Similarly, the Facebook Data Shield is a physical control panel that enables users to interact with their personal data settings on Facebook [29]. This study empathizes how difficult it can be to find the right privacy settings on Facebook and social media in general. By making the settings physical, users can toggle different data categories and see firsthand how their data is shared. The Facebook Data Shield is divided into two main layers: the core layer and the detailed layer. The core layer, which is initially visible, includes five data categories the users can interact with. The detailed layer is more complex as it has twenty-six additional data variables. The participants expressed that if the detailed layer would have been visible from the start, they would have felt overwhelmed with information. Although not all participants chose to engage with the detailed layer, the system still empowered them to take control of their data. By making data management more transparent and accessible, the Facebook Data Shield enhances users’ confidence in understanding and managing their personal information [29].

Each of these studies showcase the potential of tangible design to bridge the gap between abstract data collection and user awareness. By turning data into something physical, these studies exemplify the potential of data physicalization to make complex and often invisible processes more accessible and understandable.

3 Research Question

Making data tangible has shown to enhance its visibility and engagement, challenging individuals to confront their digital privacy practices. Data physicalization has continuously proven to be effective in helping users understand data, making abstract concepts more concrete and understandable, which enables more informed decisions. Encouraging self-reflection is particularly important in the context of cookies, where users often face a disconnect between intent and reality. By proposing a tangible approach, we can address the different challenges users face due to privacy fatigue and the privacy paradox. With this in mind, we have formulated the following research question:

How can a physical representation of cookie data make users reflect on their data privacy in a digital environment?

4 Method

In this section, we describe the process that led to our final solution. We begin by presenting the results of the initial interviews, followed by an outline of our ideation process and the methods used to explore early prototype concepts.

4.1 Initial Interviews

To gain a deeper understanding of our problem area and what a possible solution could be, we conducted user research to explore people’s perceptions and behaviors regarding cookies. We conducted five semi-structured interviews with individuals aged 25–33 using an interview guide (Appendix L). Using a semi-structured approach allowed the interview participants to freely share their perspectives and experiences while also enabling exploration on specific issues identified through our problem area [5]. The participants acknowledged that personal data is constantly being gathered, even if they do not always understand the full extent of its use.

However, none of the participants felt they had real control over their data. Participant C described this concern, saying "*I don't feel like I have control over it. I just think they take the data anyway.*" (Appendix C). Despite feeling a lack of control, some participants mentioned that they try to limit tracking by accepting only necessary cookies or rejecting them when possible. This behavior can also be linked to privacy fatigue [7]. While some participants attempt to limit tracking by adjusting cookie settings, they recognize that complete control is rarely an option. Furthermore, the interviews revealed a recurring contradiction. While the participants expressed concerns about privacy and data tracking, their actions often contradicts these concerns, supporting the privacy paradox [30]. For example, some participants insisted they always reject cookies, however they later admitted to accepting them frequently. Participant A stated, "*I always reject cookies [...]*", but later acknowledged, "*[...] but I think there are some sites that you use every day, where you've probably clicked accept.*" (Appendix A). These inconsistencies reveal the complex relationship users have with online privacy. Although users want more control and transparency, the constant need to make privacy decisions can lead to disengagement. As a result, they often accept terms they do not fully agree with, not just out of convenience, but because they grow tired of managing their privacy settings.

4.2 Ideation

Through the ideation process we followed the Design Thinking Process, an iterative approach consisting of five stages: *empathize*, *define*, *ideate*, *prototype* and *test* [12]. At the start of the project we conducted desk research and initial user interviews to get a better understanding of the problem. As we analyzed our findings, we iterated between the empathize and define stages to refine our understanding and identify the key problems. This led to the formulation of our research question [12]. Building on this process we began exploring potential solutions. We used *sketching* and *dot voting* to explore various ways to encourage users to reflect over their data privacy. First, we sketched different concepts to quickly visualize initial ideas [32]. Then we used dot voting to select the idea that resonated most with the project group [32]. The chosen concept developed into DataCrumb (Figure 4).

5 A Tangible Approach to Cookie Data

DataCrumb is a physical prototype that emulates a traditional cookie, but serves as a tangible representation of cookie data. Through DataCrumb we use data physicalization to make the abstract concept of data easier easier for users to grasp. The aim is to encourage them to reflect on their data privacy and make them reconsider before blindly accepting cookies. The prototype consists of three components: LED lights, a display screen, and a buzzer.

The decision to make the prototype look like a cookie was made due to the fact that cookies have become a symbol of data collection, often shown on cookie banners. By using a cookie as a part of our prototype, we highlight this familiar metaphor and turn it into a physical reminder of data being shared.

To visualize the flow of data queries across servers, DataCrumb has thirteen embedded LED lights that represent a specific platform or service. These platforms and services were selected based on the most used in Denmark, as well as an analysis of our own internet usage. By using Pi-hole for a seventy-two-hour period, we were able to gather our own data and compare it to the top one hundred most used websites in Denmark [1]. Streaming services were grouped into one category, as people often use different platforms. This approach helped us capture overall streaming behavior. Furthermore, we included a question mark to capture all other platforms and services.

The lights are activated when a DNS query is detected, indicating a data exchange between a device and external servers. The decision to use lights as a representation of data was made because they can easily capture the user's attention. Additionally, they provide a quick and intuitive way to visualize change in real time. Each LED on the cookie has a different color and is positioned beneath translucent M&M-shaped covers with logos, which enables users to easily identify which platform or service is receiving data.

In addition to the lights, the prototype is equipped with a buzzer. This buzzer is programmed to make a sound each time a DNS query is blocked, indicating that a tracking cookie has been prevented from transmitting data. The buzzer has two distinct sounds: a long beep and a short beep, with the long beep indicating multiple blocked queries in a five-second time frame. The buzzer was included to alert users of data being shared with third-party partners, ensuring they notice even if they are not close to the prototype.

Figure 4. DataCrumb: A physical prototype

The prototype includes a display screen that showcases the total number of queries, as well as all blocked queries. This is to give the user an overview of the amount of data being transmitted between servers.

In this project, DataCrumb is not intended as a direct solution for a specific problem but rather as a reflective and explorative tool within a Research through Design (RtD) approach. To some extent, accepting cookies and personal data collection is a requirement for navigating digital spaces, making it something that can not be avoided. Therefore, it can be seen as a "wicked problem", i.e. a problem with no clear solution [27]. With DataCrumb we align with Zimmerman et. al's (2007) argument that design artifacts can "*transform the world from its current state to a preferred state*" [34], as we aim to make users reflect and make more informed decisions regarding their online privacy. Furthermore, Gaver (2012) argues that the value of RtD is not in solving problems, but in creating artifacts that prompt reflection and new questions. The goal of this project is not to instruct users on what to do, but to make them think critically when presented with cookie acceptance banners [9].

6 Implementation

To implement DataCrumb, we followed a structured and iterative approach to ensure efficient workflow and the development of a functional prototype. We used scrum principles to guide our development process [28], dividing the work into four one-week sprints, each with a specific goal based on our prioritized product backlog. At the beginning of each sprint, we defined a sprint backlog and used Trello to plan and organize the tasks. Additionally, we used daily scrums to help us align our individual responsibilities. Weekly retrospectives gave us the opportunity to reflect on what worked well and identify areas for improvement moving forward. One of the main focuses from the beginning was to build a minimal viable product (MVP). We prioritized core functionalities such as handling incoming queries, activating the LED lights, and triggering the buzzer. Furthermore, we used pair programming and continuously tested parts of the prototype to ensure progress. Throughout the process, we used version control to keep our code organized and ensure smooth collaboration. This helped us maintain flexibility and a shared direction during the development of our prototype.

6.1 Technical setup

We built the prototype using Raspberry Pi configured with Pi-hole. The Raspberry Pi provided an accessible and portable hardware base, which made it easy to deploy the system in different environments for testing and observation. The Raspberry Pi also made it possible to connect the display screen, LED lights and buzzer to the setup (Figure 5). Pi-hole functions as a DNS sinkhole, logging all outgoing DNS queries [23]. When setting up the system, the Raspberry Pi is connected to the local network, and it's IP-address is used to configure the DNS settings on the user's devices. This simple setup enables us to monitor DNS traffic on a local network, to capture which domains were contacted as users browse the internet. Each time a domain is contacted, a LED on the prototype lights up indicating a data transmission. The thirteen lights represent the following platforms and services: *Google*, *Microsoft*, *Apple*, *Facebook*, *YouTube*, *Instagram*, *TikTok*, *Snapchat*, *LinkedIn*, *Spotify*, *Streaming*, *OpenAi* and *Other*. Furthermore, we use Pi-hole as network-wide protection, for blocking third-party tracking [24]. These blocked queries serve as a trigger for activating the buzzer sound.

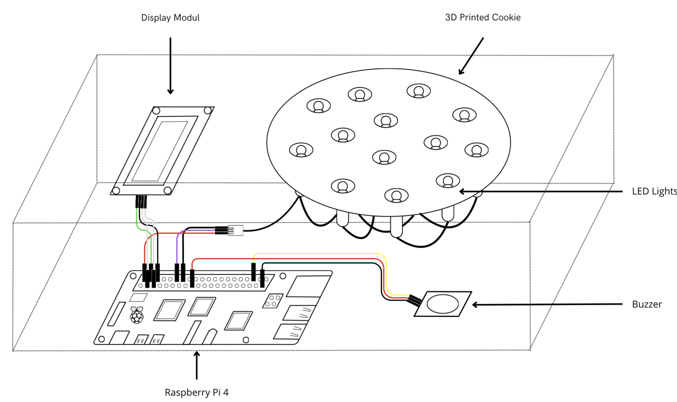

Figure 5. System setup

6.2 Pseudo code

The following sections will explain the functionalities of DataCrumb using pseudocode, to provide a clear and straightforward overview of the system's logic and structure. During the development of the prototype, we chose to separate the functionalities of the three components, to ensure modularity, easier debugging, and improved maintainability. Additionally, we created a main file that runs the overall program. This main file connects to the individual functionalities and coordinates their interactions.

6.2.1 Display. The display shows two counts: total number of queries and number of blocked queries (Figure 6). The display screen is turned on when the system starts running, and starts by setting each count to zero. Each time a new query is detected though Pi-hole, the counts get updated.

```
1.  Turn on the display
2.  Set total queries to 0 and blocked queries to 0
3.  Set display to show current number of total and
    blocked queries
4.  For each new query
5.      Increase the total count by 1
6.      If the query is blocked
7.          Increase the blocked count by 1
8.      Update the display with the new counts
```

Figure 6. Pseudocode: Display

6.2.2 Buzzer. Figure 7 shows the buzzer code, which keeps track of all the blocked queries and activates the buzzer. It starts by setting a count of blocked queries to zero. By looping through each incoming query it checks their status, and updates the count accordingly. If the status is equal to GRAVITY, it means that the query is blocked by Pi-hole. For each blocked query detected, the count on the display is updated. To clearly indicate when multiple blocked queries are detected, we use two distinct buzzer sounds: a short beep for single queries and a long beep for multiple queries. By applying a five second interval, we can detect instances where multiple queries are being blocked simultaneously and update the count for each blocked query. This is used to determine which sound is triggered based on the number of counts after each interval. These intervals are repeated while the system is running, and after each interval the count is reset to zero.

```
1.  Set count to 0
2.  Loop through incoming queries
3.      If the query has the status GRAVITY
4.          Increase the blocked count with 1
5.          Run display code to update the blocked count

6.  Set a time interval for 5 seconds
7.      For each 5 second interval
8.          Check count
9.              If count is equal to 5 or more
10.                 Activate long beep
11.             If count is equal to 0
12.                 Don't activate the buzzer
13.             Otherwise
14.                 Activate short beep
15.             Reset count to 0
16.         Start a new interval
```

Figure 7. Pseudocode: Buzzer

6.2.3 LED lights. As shown in figure 8, every time a new domain is detected with Pi-hole, the associated light is turned on for twenty seconds. If the same light is triggered repeatedly, it will remain lit until each activation has completed their twenty second duration. There are thirteen different LED lights, one for each platform and service selected to be represented on the prototype.

```
1.  When a domain is detected
2.      Turn on the associated LED with a specific color
3.      After 20 seconds, turn off the light
```

Figure 8. Pseudocode: LED lights

6.2.4 Main. Figure 9 shows the pseudocode for the main file that runs the code for DataCrumb. It begins by connecting to Pi-hole to gain access to incoming DNS queries on the network. After that the system enters a loop where it continuously checks for new DNS queries from Pi-hole. If no new queries are found, the system starts the loop from the top. This ensures that it keeps monitoring for activity without performing unnecessary actions. The system first checks whether the query is of type HTTPS (equivalent to A or AAAA domain query), to ensure that only queries potentially involving cookies are processed. If the query is HTTPS, the system proceeds to check the domain name. When checking the domain name, the system runs through a loop of the chosen platforms. For each match it activates the corresponding LED light, to visually indicate which platform the query is associated with (Figure 8). After identifying the domain, it checks whether the domain is blocked, by running the code for the buzzer (Figure 7). Finally, the query is counted in the total amount of queries count, and logged for the display (Figure 6).

```
1.  Start background task to handle buzzer time interval
2.  Connect to Pi-hole
3.  Loop continuously
4.      Wait for 1 second
5.      Request new DNS queries from Pi-hole
6.      If no new queries are found
7.          Starts the loop from the top
8.      For each new detected query
9.          If the query is of type HTTPS
10.             Check which platform the domain belongs to
11.                 Run light code for this specific domain
12.             Run buzzer code to check for blocked query
13.             Run display code to update the total amount
```

Figure 9. Pseudocode: Main

7 Evaluation & Insights

After developing DataCrumb, we wanted to test the prototype. At this stage of the Design Thinking Process, the objective is to gather new insights that help deepen our understanding of the problem and identify opportunities to refine the prototype [12]. The test was conducted with three participants, with the intent to explore whether the prototype would encourage reflection on their use of cookies, as well as impact their awareness regarding the sharing of cookie data and their own data privacy. The test comprised three parts, each serving a specific purpose: a pre-interview, an in-home test session, and a post-interview (Appendix M). The pre-interview was included to get an understanding of the test participants' thoughts on cookies before testing the prototype. The in-home test consisted of a three-day period where the prototype was placed in each person's home. All of the participants' devices were connected to Pi-hole, and the prototype would react with lights and sounds, without any interaction from the participants. The placement of the prototype was made based on where the participants usually used their devices, so that it was visible when the participants accessed the internet and used their mobile phone, computer and television. The prototype was placed in the bedroom and living area for participants 1 and 3, and in the living room for participant 2. After the in-home test, we conducted a post-interview. The purpose with this interview was to get insights on the participants' experience with having the prototype in their home, and whether it had any influence on their attitudes or behaviors regarding cookies and data privacy.

7.1 Findings

To organize the user data gathered through the tests, we created an Affinity Diagram based on a brainstorming session [13]. By writing out statements made by the participants on post it's, and grouping them into categories by similarity, we were able to identify common experiences and issues (Figure 10). The categories identified through this process were the following: *Cookie Practices, Helplessness, Intent vs Reality, Awareness of Data Collection,* and *Impressions.*

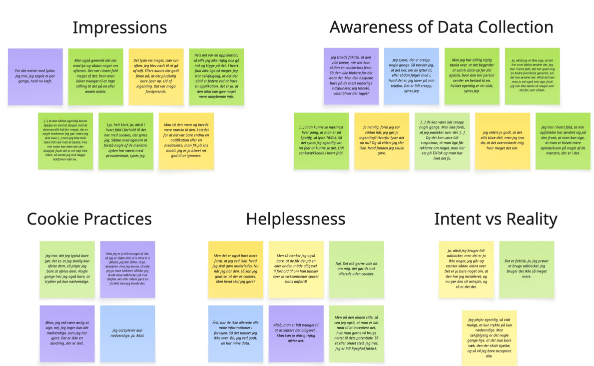

Figure 10. Affinity diagram

7.1.1 Cookie Practices. During the pre-interview the participants were asked about their interactions with cookie banners. All three participants stated that when presented with the choice, they made an effort to only accept necessary cookies. After the three-day test period, their opinion on the matter was unchanged, with P2 saying *"Uh, I have to be honest and say no. I only choose what is necessary, as I have always done. It's not a change that has happened."* (Appendix J). This suggests that the prototype had little impact on the participants' future interactions with cookie banners.

7.1.2 Helplessness & Intent vs Reality. Even though the participants said that they only selected necessary cookies, they contradicted themselves by stating that they often accept all cookies, in order to make the cookie banner disappear. The statement *"I usually, as much as possible, only click on "necessary only." But of course, sometimes I just want that damn banner to go away, so I end up accepting all."*(Appendix F) highlights this contradiction. Furthermore, the participants express a desire to make active changes and use tools that help them protect their data privacy, but also say that they do not really think about it in their daily lives. This behavior reflects the privacy paradox described in Section 2.2. The failure to safeguard their data privacy stems from a lack of knowledge on how to do so. The participants state that they do not know what to do differently, and that they feel forced to share their data through cookies. Additionally, there seems to be a sense of resignation and helplessness, as they believe that different platforms and companies already have access to their data. In the pre-interview P2 states *"Oh, don't they already have all my information anyway? So I don't really think about it. Uh, I know they have my data."* (Appendix G). This feeling of lack of control aligns with the challenges faced by users experiencing privacy fatigue, as explained in Section 2.2. Overall, the inconsistencies in participants' behavior, their attitudes regarding data privacy, as well as their state of mind corresponds with the findings from the initial interviews in Section 4.1.

7.1.3 Awareness of Data Collection. In the pre-interviews, all the participants expressed concerns about how companies use their data, describing it both as suspicious and creepy. P2 stated *"I think it's creepy sometimes. I start wondering if they're listening or somehow keeping track of what I'm doing on my phone. I find that a bit creepy."*(Appendix G). This highlights how the participants were unsettled by the idea that companies can monitor their online behavior, particularly through tracking cookies. P3 also mentioned finding it strange when they received advertisements for something they had viewed or liked on another platform, which reinforces their sense of being constantly monitored (Appendix H).

DataCrumb contributed to an increased awareness of data collection among the participants. Although they were aware that their data was being collected and shared, they were still shocked to see just how much was being collected in real time. The participants were surprised that data was collected even when they were not actively using their devices. P1 said, *"Exactly, because I was like, I'm not even doing anything? Why is it lighting up now? And then I had no idea what the hell I was supposed to do."* (Appendix I). This reaction shows how the prototype helped reveal the hidden processes of data tracking, which is typically something that goes unnoticed. It gave the participants a clearer understanding of how much information is collected passively, which led to increased concern and reflection about their digital privacy. This contrast between what the participants thought they knew and what they observed also highlights the abstract nature of cookies and data privacy. The prototype helped visualize the otherwise hidden flow of data between platforms, which sparked critical reflection about online behavior. P3 mentioned that they became more aware of this pattern of data sharing between platforms: *"[...] you could see almost every time you were on Spotify, TikTok would light up. So I actually thought it was pretty cool to be able to see that. Definitely thought-provoking."*(Appendix K). Additionally, they said, *"I definitely think my perception has changed in that regard, you could say I've become more aware of some of the patterns involved."* (Appendix K). These statements highlight how DataCrumb made invisible data processes more visible, which led to a deeper understanding and increased awareness of digital data sharing.

7.1.4 Impressions. Considering the individual components of the prototype: lights, buzzer and display, different aspects stood out to the participants. However, one point was consistent: All participants initially found the beep sound provocative, as it forced them to acknowledge the presence of the prototype. Over time, P1 and P3 reported that they became accustomed to it.

For P1, the display and lights made the strongest impressions. They explained seeing the amount of shared data visualized on the display was eye-opening, as it showed how much data was actually shared. The lights, in particular, had an effect on their behavior. They chose to disconnect from the internet at night, so they could sleep, since the prototype continuously lit up their bedroom. In contrast, for P2, the sound was the most effective in drawing attention to the prototype, as they found it impossible to ignore. When reflecting on their experience, they said *" [...] I think I said 'shut up' a couple of times. I felt it beeped a few too many times [...]"* (Appendix J) when it came to the effect of the sound. P3 found both the lights and sound to be thought provoking. These two components pushed them to confront their own digital habits, particularly in the evening. As a result of the beeping sound, they deliberately used their phone less, to minimize the noise. When asked about what made the biggest impression, P3 said: *"The Lights, definitely. Yeah, at least when it comes to the cookies, that's what I think. Like understanding some of the patterns. The sound has been the most provocative, I think"* (Appendix K). Together, these responses highlight DataCrumb's different components and how it has made the participants reflect on their understanding of cookies. Overall, having the prototype in their home led the participants to reflect on their use of digital devices.

During the post-interview, two participants discussed the advantage of having a physical prototype compared to a digital one. P1 noted *"You noticed it more and paid more attention to it. Instead of it just being another notification or message you get on your phone. I have become pretty good at ignoring those"* (Appendix I). This demonstrates that, unlike digital notifications, the physical presence of the prototype made it harder to ignore, and encouraged deeper reflection on data privacy.

The lights effectively captured the participants attention, while visualizing data sharing, as indented. Additionally, the buzzer served as a provocative element. The display provided a general overview of the number of queries over the three day period, along with the amount of blocked queries. While the display was found useful by two participants, P3 expressed a desire for additional information to be shown. This aspect will be explored further in Section 8.

Figure 11. An example of test setup

7.1.5 Key Takeaways. The findings indicated that DataCrumb did not lead to a direct change in the participants behavior regarding cookie banner interactions. However, the prototype effectively increased their awareness of data collection and promoted reflections on their data privacy. The challenges identified during the test sessions aligned with those described in Section 2.2, as well as the findings from the initial interviews, which indicated that users experience both privacy fatigue and the privacy paradox. Making the prototype physical and visualizing the data using lights made the underlying processes of data sharing more tangible for the participants, which in turn encouraged self-reflection. Overall the different components of DataCrumb made impressions on the participants, and helped them connect abstract data concepts to personal experiences.

8 Discussion

This section discusses the physical setup of the prototype during the test sessions and examines participants' responses to the display component. Furthermore, it considers the potential of a more provocative approach and explores how this might influence the outcomes. Additionally, we address the limitations of using Pi-hole as part of the setup, and discuss possibilities for future iterations of the project.

8.1 Prototype Placement and Participant Interaction

During the setup of the prototype in the participants' homes, the prototype was placed in the area where each participant used their devices most frequently. One participant lived in a two-bedroom apartment, while the other two lived in studios. As a result, the placement of the prototype varied (see Section 7). This variation in placement raises the question of whether the position influenced how participants experienced and interacted with DataCrumb.

During the testing of the prototype, it became clear which components made the strongest impression on the participants. For example, P2 who had the prototype placed in the living room, responded primarily to the sound rather than the lights (Appendix J). The sound was mostly noticeable in the evenings when the participant was in their bedroom. In contrast, as P1 and P3 live in a studio apartment, with only one main living area, the prototype remained visible at all times, which appeared to increase their awareness of the lights (Appendix I, K). This could indicate that placement of the prototype can influence which aspects of the prototype were most dominant for the participants' experience. If the placement of the prototype had been the same across all participants, there may have been more of a standardized baseline for comparison. However, similar placement might also result in positioning the prototype in rooms where the participants used their devices less frequently, potentially reducing the prototype's overall impact. This suggests that the placement could be an important consideration for future research. Additionally, exploring the effects of deploying multiple prototypes within the home could be valuable, particularly to determine whether this reduces the possibility of the participants overlooking the different components of the prototype.

The test session also provided insights into participants' experience with the display. The findings varied, as two participants indicated that the display had minimal impact: one initially noticed it but quickly lost interest, while the other did not mention it at all during the final interview (Appendix I, J, K). In contrast, the third participant found the display engaging, stating that seeing the actual numbers representing shared data had a stronger effect. These varied responses to the display may be due to the short testing period. The prototype was tested over a period of three days, which may not have been sufficient time for the participants to fully engage with or reflect on the value of the display. It is possible that, over a longer period, participants would become more aware and engaged with the display as the numbers increase. Further exploration could benefit from extending the test duration to observe whether long-term exposure could enhance the impact of the display component and encourage deeper reflection on data sharing.

8.2 A Provocative Approach to Cookie Data

As previously mentioned, the purpose of DataCrumb is to prompt reflection and raise awareness. While provocation was not a focus of the prototype, the tests revealed that the buzzer acted as a provocative element. The intended function of the buzzer is to serve as an alert each time queries are blocked due to tracking cookies. Since queries were frequently blocked during the test period, the sound became a disruptive reminder and interruption in their usual routines. The unexpected provocation led to some changes in behavior among the test participants, such as limiting their phone usage to reduce the number of times the buzzer was triggered. This suggests that provocative elements have the potential to influence user behavior. By incorporating a provocative design approach we could possibly have changed the participants' interactions with cookie banners, which DataCrumb does not achieve in its current state. As a design approach, provocation can be used to challenge user expectations and disrupt habitual practices to provoke reflection and encourage alternative behaviors [2, 25]. Although incorporating provocation holds potential when the intention is to change behavior, it comes with challenges, as provocative elements can lead to disengagement rather than reflection.

8.3 Limitations

To gather the data required for the prototype to function as intended, we used Pi-hole. Through Pi-hole we are able to capture all HTTPS queries made on a network while using the prototype, but we cannot be certain that each of these queries contains cookie data. This means that while our prototype visualizes which platforms and services gather data, it does not specifically show cookie-related traffic. In that sense, Pi-hole presents a limitation in relation to our research focus on cookie activity. Despite this limitation, Pi-hole serves as a way to effectively gather data in real-time without building a custom solution from scratch. Furthermore, it operates at the network level and can be connected to multiple devices on a home network seamlessly, which makes it practical and efficient for supporting the functionality of DataCrumb.

8.4 Future Work

The physical nature of the prototype had the intended impact on the test participant. DataCrumb helped make the abstract concept of data more tangible, as well as increased the participants' awareness on data sharing and made them reflect on their own data privacy. The tests revealed that providing additional information would be helpful to offer a deeper understanding and more contextualized details about the data being shared. This information could be integrated into an application that would be used alongside the physical prototype. By creating a digital graph of the data visualized on the prototype, we would effectively show how data moves between different platforms, and which platforms are connected by this data. In addition to showing the data flow, the application could provide a more detailed explanation of the thirteen chosen platforms. This could be especially beneficial for the "Other" category, as this encompasses multiple different platforms and services. The possible development of an application for future iterations of DataCrumb could also make the prototype easier to customize for the specific user, as additional categories could be added besides the chosen thirteen platforms. Additional categories could also be included on the physical prototype. Even though this is the case, adding categories and lights could make the amount of data overwhelming, which would undermine the goal of making data tangible and easily understandable. Overall, balancing the level of detail and simplicity will be crucial in future iterations to ensure that DataCrumb remains both engaging and accessible, enabling users to better understand and reflect on their data privacy.

9 Conclusion

This project has explored how a physical representation of cookie data can help encourage reflection on data privacy in a digital environment. One of the core challenges with cookie data is its abstract nature and the difficulty users face in understanding its collection. This is further complicated by deceptive design practices, privacy fatigue and the privacy paradox. As highlighted in the Related Work section, these factors can hinder users' ability to meaningfully engage with decisions about their own digital privacy. To address this problem, we looked into studies on data physicalization and tangible designs. These approaches showed how transforming abstract data into comprehensible designs can promote self-reflection and awareness. Based on these insights, we formulated the following research question:

How can a physical representation of cookie data make users reflect on their data privacy in a digital environment?

Our solution was the design and development of DataCrumb: a physical prototype that visualizes how cookie data gets shared through lights, sound and a display screen. The prototype does not aim to solve the issue of online tracking through cookies, but functions as a tool for awareness and reflection. Our tests revealed a greater awareness of data collection and its underlying processes. The participants were surprised by the amount of data shared, even when they didn't use their devices actively, which led to increased concerns about their data privacy. Additionally, the physical aspect of the prototype made it hard for the participants to ignore, thereby prompting reflection on how their data is collected and used. These findings support the potential of tangible and physical designs in addressing digital privacy concerns.

References

[1] Ahrefs. [n. d.]. Top Websites in Denmark - May 2025. https://ahrefs.com/websites/denmark. Accessed: 2025-05-27.

[2] Shaowen Bardzell, Jeffrey Bardzell, Jodi Forlizzi, John Zimmerman, and John Antanitis. 2012. Critical design and critical theory: the challenge of designing for provocation. In *Proceedings of the Designing Interactive Systems Conference* (Newcastle Upon Tyne, United Kingdom) *(DIS '12)*. Association for Computing Machinery, New York, NY, USA, 288–297. doi:10.1145/2317956.2318001

[3] Benjamin Maximilian Berens, Heike Dietmann, Chiara Krisam, Oksana Kulyk, and Melanie Volkamer. 2022. Cookie Disclaimers: Impact of Design and Users' Attitude. In *Proceedings of the 17th International Conference on Availability, Reliability and Security* (Vienna, Austria) *(ARES '22)*. Association for Computing Machinery, New York, NY, USA, Article 12, 20 pages. doi:10.1145/3538969.3539008

[4] F. Botros, C. Perin, B.A. Aseniero, and S. Carpendale. 2016. Go and Grow: Mapping Personal Data to a Living Plant. In *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16)*. Association for Computing Machinery, New York, NY, USA, 112–119. doi:10.1145/2909132.2909267

[5] Alan Bryman. 2015. *Social Research Methods* (5 ed.). Oxford University Press. 201–202 pages.

[6] Geni Bushati, Sven Carsten Rasmusen, Anelia Kurteva, Anurag Vats, Petraq Nako, and Anna Fensel. 2024. What is in your cookie box? Explaining ingredients of web cookies with knowledge graphs. *Semantic Web* 15, 5 (2024), 1593–1609. doi:10.3233/SW-233435 arXiv:https://journals.sagepub.com/doi/pdf/10.3233/SW-233435

[7] Hanbyul Choi, Jonghwa Park, and Yoonhyuk Jung. 2018. The role of privacy fatigue in online privacy behavior. *Computers in Human Behavior* 81 (2018), 42–51. doi:10.1016/j.chb.2017.12.001

[8] Pierre Dragicevic, Yvonne Jansen, and Andrew Vande Moere. 2021. Data Physicalization. In *Springer Handbook of Human Computer Interaction*, Jean Vanderdonckt, Philippe Palanque, and Marco Winckler (Eds.). Springer, Cham, Chapter 94. doi:10.1007/978-3-319-27648-9_94-1

[9] William Gaver. 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 937–946. doi:10.1145/2207676.2208538

[10] Alejandra Gomez Ortega, Renee Noortman, Jacky Bourgeois, and Gerd Kortuem. 2024. Dataslip: Into the Present and Future(s) of Personal Data. In *Proceedings of the Eighteenth International Conference on Tangible, Embedded, and Embodied Interaction* (Cork, Ireland) *(TEI '24)*. Association for Computing Machinery, New York, NY, USA, Article 40, 14 pages. doi:10.1145/3623509.3633388

[11] Paul Graßl, Hanna Schraffenberger, Frederik Zuiderveen Borgesius, and Moniek Buijzen. 2021. Dark and Bright Patterns in Cookie Consent Requests. *Journal of Digital Social Research* 3, 1 (February 2021), 1–38. doi:10.33621/jdsr.v3i1.54

[12] Interaction Design Foundation. 2016. Design thinking (DT). Accessed 22-05-2025 https://www.interaction-design.org/literature/topics/design-thinking.

[13] Interaction Design Foundation. 2024. *Affinity Diagrams*. https://www.interaction-design.org/literature/topics/affinity-diagrams Accessed: 2025-05-28.

[14] Hiroshi Ishii and Brygg Ullmer. 1997. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '97)*. Association for Computing Machinery, New York, NY, USA, 234–241. doi:10.1145/258549.258715

[15] Yvonne Jansen, Pierre Dragicevic, Petra Isenberg, Jason Alexander, Abhijit Karnik, Johan Kildal, Sriram Subramanian, and Kasper Hornbæk. 2015. Opportunities and Challenges for Data Physicalization. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, Seoul, Republic of Korea, 3227–3236. doi:10.1145/2702123.2702180

[16] Rohit Ashok Khot, Florian 'Floyd' Mueller, and Larissa Hjorth. 2013. SweatAtoms: Materializing Physical Activity. In *Proceedings of the 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death* (Melbourne, Australia) *(IE '13)*. Association for Computing Machinery, New York, NY, USA, Article 4, 7 pages. doi:10.1145/2513002.2513012

[17] Richie Koch. 2024. Cookies, the GDPR, and the ePrivacy Directive. https://gdpr.eu/cookies/ Accessed: 2025-06-10.

[18] Michael Kretschmer, Jan Pennekamp, and Klaus Wehrle. 2021. Cookie Banners and Privacy Policies: Measuring the Impact of the GDPR on the Web. *ACM Trans. Web* 15, 4, Article 20 (July 2021), 42 pages. doi:10.1145/3466722

[19] Maxwell Lin, Shihan Lin, Helen Wu, Karen Wang, and Xiaowei Yang. 2024. Browsing without Third-Party Cookies: What Do You See?. In *Proceedings of the 2024 ACM on Internet Measurement Conference* (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 130–138. doi:10.1145/3646547.3689014

[20] Eryn Ma and Eleanor Birrell. 2022. Prospective Consent: The Effect of Framing on Cookie Consent Decisions. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI EA '22)*. Association for Computing Machinery, New York, NY, USA, Article 400, 6 pages. doi:10.1145/3491101.3519687

[21] Thomas Mejtoft, Nike Vejbrink Starbrink, Carla Roos Morales, Ole Norberg, Mattias Andersson, and Ulrik Söderström. 2023. Cookies and Trust: Trust in organizations and the design of cookie consent prompts. In *Proceedings of the European Conference on Cognitive Ergonomics 2023* (Swansea, United Kingdom) *(ECCE '23)*. Association for Computing Machinery, New York, NY, USA, Article 18, 6 pages. doi:10.1145/3605655.3605668

[22] Bayan Al Muhander, Omer Rana, Nalin Arachchilage, and Charith Perera. 2022. Demo Abstract: PrivacyCube: A Tangible Device for Improving Privacy Awareness in IoT. In *Proceedings of the 2022 IEEE/ACM Seventh International Conference on Internet-of-Things Design and Implementation (IoTDI)*. IEEE, New York, NY, USA, 109–110. doi:10.1109/IoTDI54339.2022.00024

[23] Pi-hole. 2022. *Pi-hole Documentation*. https://docs.pi-hole.net/ Accessed: 2025-05-22.

[24] Pi-hole. 2024. *Pi-hole*. https://pi-hole.net/ Accessed: 2025-05-22.

[25] Dimitrios Raptis, Rikke Hagensby Jensen, Jesper Kjeldskov, and Mikael B. Skov. 2017. Aesthetic, Functional and Conceptual Provocation in Research Through Design. In *Proceedings of the 2017 Conference on Designing Interactive Systems* (Edinburgh, United Kingdom) *(DIS '17)*. Association for Computing Machinery, New York, NY, USA, 29–41. doi:10.1145/3064663.3064739

[26] Ali Rasaii, Shivani Singh, Devashish Gosain, and Oliver Gasser. 2023. Exploring the Cookieverse: A Multi-Perspective Analysis of Web Cookies. In *Passive and Active Measurement: 24th International Conference, PAM 2023, Virtual Event, March 21–23, 2023, Proceedings*. Springer-Verlag, Berlin, Heidelberg, 623–651. doi:10.1007/978-3-031-28486-1_26

[27] Horst W. J. Rittel and Melvin M. Webber. 1973. Dilemmas in a General Theory of Planning. *Policy Sciences* 4, 2 (1973), 155–169. doi:10.1007/BF01405730

[28] Ken Schwaber and Jeff Sutherland. 2020. The Scrum Guide – The Definitive Guide to Scrum: The Rules of the Game. Accessed: 2025-06-10.

[29] Jules Sinsel, Anniek Jansen, and Sara Colombo. 2023. Facebook Data Shield: Increasing Awareness and Control over Data used by Newsfeed-Generating Algorithms. In *Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction* (Warsaw, Poland) *(TEI '23)*. Association for Computing Machinery, New York, NY, USA, Article 46, 6 pages. doi:10.1145/3569009.3573116

[30] Daniel J. Solove. 2020. The myth of the privacy paradox. *SSRN Electronic Journal* (Jan. 2020). doi:10.2139/ssrn.3536265

[31] Jie Tang, Umair Akram, and Wenjing Shi. 2020. Why people need privacy? The role of privacy fatigue in app users' intention to disclose privacy: based on personality traits. *Journal of Enterprise Information Management* ahead-of-print (08 2020). doi:10.1108/JEIM-03-2020-0088

[32] TISDD Method Library. n.d.. This Is Service Design Doing. Accessed 22-05-2025. https://www.thisisservicedesigndoing.com/methods.

[33] Christine Utz, Martin Degeling, Sascha Fahl, Florian Schaub, and Thorsten Holz. 2019. (Un)informed Consent: Studying GDPR Consent Notices in the Field. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security* (London, United Kingdom) *(CCS '19)*. Association for Computing Machinery, New York, NY, USA, 973–990. doi:10.1145/3319535.3354212

[34] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 493–502. doi:10.1145/1240624.1240704