Design Project - The Citizenlab

Building a health platform for citizen science

Mark Boom (s2552469)
Silas de Graaf (s2220032)
Dawid Kulikowski (s2472910)
Jesse Snoijer (s2572362)
Joep Vorage (s2172968)

April 2023

Abstract

Citizen science is all about practicing science in collaboration with the general public. This paper constitutes a report of a ten-week project by students in collaboration with the Citizenlab at the University of Twente. The goal of the project was to specify, design, implement and present an online platform for citizen science. The project put an emphasis on specific needs for researchers such as a well-designed flow for dynamic informed consent as well as a module for streamlined communication with participants that still safeguards their anonymity. Other requirements like complete control over data and transparency are covered as well. Citizen science becomes more useful when participants can meaningfully use their own data, which is why the platform includes features to empower insights for citizens. Overall, privacy and security formed important factors within the project. The need for general applicability of the platform also implies that an accessible and easy to understand user interface is of vital importance. How do users - both participants and researchers - want to experience conducting scientific studies in collaboration with each other? What digital facilities are required for such an endeavour to succeed? Literature research into best practices was performed, and based on requirements gathered from researchers through interviews were transformed into a design which was ultimately implemented and tested. Proposed future work includes making the platform more robust and conducting further user testing. Throughout the report the built platform is described, in hopes it can eventually serve as a prototype for further research.

Preface

Over the past ten weeks, we participated in the Design Project and completed a full design cycle for a software system. Based on an initial proposal by the Citizenlab - a research group at the University of Twente - requirements were elicited for what would be a health platform for citizen science. With a goal of building the core of such a software system, requirements were transformed into a design and an implementation followed. This report will explain the design and the choices behind it as well as reflect on the process.

Acknowledgements

We would like to thank our supervisor, Lorenzo Gatti[1], for his valuable feedback and guidance during the project. Furthermore, we would like to express our gratitude to Ria Wolkorte and Lieke Heesink [2] for providing the project proposal and actively participating in our design process. We thank the jury of the programme for naming us best project of the year.

[1] Dr. Gatti is an assistant professor at the Human Media Interaction (HMI) group of the University of Twente

[2] Dr. Wolkorte and Dr. Heesink are assistant professors at the University of Twente and connected to the Citizenlab

Table of Contents

1 Introduction

1.1 The Citizenlab

The Citizenlab is an organization which performs research in active collaboration with the general public ("citizens"). This research is often focused on health-related topics. For example, the Citizenlab is currently focusing on a study concerning rheumatoid arthritis. For such a study, citizens are asked to fill out a daily survey with several questions that are relevant to its particular subject. At the end of the research period, the data is collected, transformed and analyzed to extract meaningful insights and conclusions.

While the primary goal of the Citizenlab is doing research, it greatly concerns itself with the needs and wishes of these citizens as well. They are able to voice which aspects of research are most relevant and studies are designed in close collaboration with them.

1.2 The health platform

During the Design Project, we were tasked with designing and creating a health platform for the Citizenlab. The health platform is an online web application where researchers of the Citizenlab are able to conduct surveys for their research and citizens are able to fill out surveys as well as view their own data. It is meant to be a one-stop environment where data can be collected, stored, viewed and exported. Such an application comes with many requirements in terms of functionality, security, user friendliness and correctness. All of this will be further detailed in the requirements section of this report. To read more about the problem statement, please refer to the problem section of our original project proposal.[3]

1.3 Summary of proposal

We started the Design Project by having an initial interview with the client. The information gathered through this interview was then used to write a project proposal. The project proposal contains an investigation of the current problem as well as a list of requirements for the solution. This is accompanied by a plan for further research, a risk assessment and potential timeline. Most information included is based on the initial interview with our clients.

1.3.1 Problem

The researchers of the Citizenlab wish to have their own platform to conduct surveys for their research. This platform should allow for simple survey creation with extensive customizability. The platform should provide value to citizens by giving them the ability to view and extract their own data. Its design should be intuitive and present a greater ease of use than the current solution.

1.3.2 Research

To assist us in making initial decisions about, among other things, accessibility issues we performed literature research. We mainly considered research done into the accessibility of websites for elderly people as well as people that are physically impaired in some manner. We found that people which would fit the description of a citizen follow a different reading pattern, prefer bigger buttons, like text over icons and find it difficult to scroll. We took all these things into account while designing and implementing a solution to the problem.

1.3.3 Solution

The proposed solution is a new platform which would facilitate conducting surveys which must be filled out periodically and provide a way for citizens to communicate anonymously with the researchers. It differs from current commercial solutions and the platform currently being used by our clients in the following ways:

[3] Appendix F

- The platform will facilitate creating and editing surveys as a researcher which is not possible in the current solution.
- The platform will be tailored specifically to the Citizenlab's target group, mostly elderly people, while the current solution is mainly targeted towards young people.
- The platform will allow for anonymous communication between researchers and citizens while the current solution does not have any way for researchers to contact the people participating in their surveys.

1.3.4 Risk

A vital part of the project proposal was the risk analysis. We identified a variety of risk classes ranging from privacy and security risks to planning risks.

The main risk we identified is related to the sensitive data stored by the platform. Any data leak can contain strictly confidential information ranging from a personal diary to blood test results. Furthermore, depending on how the citizens use the platform, an error is displaying the data can cause direct harm to the user by, for example, causing them to take the wrong dosage of a medication. The final risk we identified had to do with planning and failing to deliver a usable product on time or meet other deadlines.

1.4 Glossary

In the glossary below, you will find some terms which are specific to this project and may not be common knowledge. Defining these terms upfront in the process helped us communicating with the clients.

citizen (with rheumatoid arthritis), participant	user filling in surveys
client(s)	researchers from the Citizenlab who are involved in the project
health platform	the product / (web) application we create
researcher	user creating and managing surveys
(end) users	citizens & researchers
dynamic informed consent	extended "informed consent" procedure which provides the option to control consent on a very granular level and allows for adjustment of consent as often as desired [4]

2 Design and requirements

A good design of the software system was a primary objective of this project. This section will lay out different elements of that design as well as highlighting the design *choices* that led to the final product. This section is divided up into six larger themes that occur throughout the system and stem directly from the requirements.

2.1 Dynamic Informed Consent procedure

Processing user data for research purposes presents unique challenges, especially considering the fact that data collected by the platform is in most cases of a sensitive nature as users will often be tracking their own health. Before gathered data can be used for scientific research, user consent must be unambiguously established: institutes that are carrying out the research, such as a university, uphold high ethical standards that make user consent of primary concern.

[4] A more detailed explanation can be found in section 2.1 of the report.

Our clients expressed the desire to work with a *Dynamic Informed Consent* (DIC) procedure. This is a relatively new, but important procedure within citizen science. The client was especially interested in what is possible in the technical world for this since they did not know of any current survey platform that uses it. DIC inherits all characteristics from a regular informed consent procedure, but adds a couple of new principles too.

2.1.1 Characteristics of a regular informed consent procedure

Important characters of a regular informed consent procedure for research are:

- **Voluntary consent**
  User consent must be voluntary, which means that a person giving consent may not feel any pressure to do so. Such pressure can be applied directly (threatening a participant with negative consequences upon refusal to participate), but often is more subtle. Voluntary consent can be safeguarded by actively avoiding this pressure, such as explicitly stating there a no negative consequences connected to non-participation.
- **Informed consent**
  A given consent is of little value when a person does not fully grasp the consequences of their consent. At all times, they must know exactly what happens to their data and why. To safeguard the informed characteristic of the consent, ample information must be provided both while giving the consent as well as after the consent has been given.
- **Right to retract**
  Participants may retract their consent at all times, without giving any reason to do so and without negative consequences. They need to be well-informed about this and it is important the procedure to retract consent works properly and is honoured by researchers.
- **Opt-in**
  Informed consent is always opt-in, and can never be implicitly given by a user.

2.1.2 Characteristics of a Dynamic Informed Consent procedure

The DIC adds more principles to augment a regular informed consent procedure:

- **Dynamic updating of consent**
  Participants can update their consent at any time in an easy manner. This means re-specifying granting consent or retracting consent is possible at all times through a consent wizard.
- **Transparency**
  A DIC procedure is very transparent, and users can see when their data was used and why. Obviously, this transparency is limited to the disposition of the researcher, as they can choose to provide untruthful reasons for a data extraction. However, it does force researchers to state on the record why data was extracted, allowing them to be held responsible in case of misuse.

2.1.3 Limits of informed consent

The end product of scientific research will contain user data as well, albeit aggregated and anonymized. After all, the goal of data collection is to use this data in the research. Even in a DIC procedure, users cannot retract their consent for data already used in research, as this is practically impossible. There are two ways user data permanently ends up in scientific research:

- Data is used in the final product, such as a paper. It is known beforehand when this data will be extracted, and hence a user knows beforehand what the deadline is to decide if they consent to their data being used in the research.

- Data can also be extracted on a preliminary basis, for example to validate the health of the data or to present intermediate results at a conference. These moments are not established beforehand, and users hence cannot predict or adjust their consent for these extractions.

In both cases, two principles justify the permanent usage of data. First of all, data is aggregated and anonymized, so no data can be traced back directly to a single participant. This is simply a principle of conducting scientific research in an ethically sound manner. Secondly, user consent at the time of data collection counts. Users are informed beforehand of both ways their data may permanently end up in scientific research and can thus plan accordingly, even though preliminary data extractions are not planned in advance: if the given consent at that point was justifiable by adhering to the principles of the procedure, use of the data by researchers is justified.

2.1.4 Design of the DIC procedure

To meet all requirements from the client regarding the consent procedure and settings, we deliberated amongst ourselves and came up with three different areas of consent which would together cover the different needs of the health platform. After we came up with a concrete plan regarding this, we showed it to our clients who approved the design.

We spent a lot of time discussing the different forms of consent with the client. In the end, we ended up with the following forms of consent which were approved by the client:

- **Primary consent**
  This constitutes (dynamic) informed consent to use a users' data in a predefined study that has a set goal, background and start- and end dates.
- **Secondary consent**
  Researchers may want to use data that the participant is going to collect in another research that does not exist yet. By giving secondary consent, a participant grants the researchers consent to approach them once such a research opportunity arises. At that point, the user can either grant or give (primary) consent to use their (already collected) data in the new research. Hence, an important feature of secondary consent is thus that it gives researchers the opportunity to ask for primary consent of a study that does not exist yet; the secondary consent itself does not allow for the usage of user data.
- **Tertiary consent**
  Sometimes, researcher would like to communicate to participants if new research is being conducted that may be of interest to participants of the current study. This may also include other announcements that are related to, but not essential to the current survey. Consent is required to approach users with such announcements or requests. Note that just like with secondary consent, the tertiary consent does not allow for using user data in research in any way.

All types of consent adhere to the principles of DIC mentioned in the previous section.

A schematic overview of the three types of consent together with required flows is provided in Appendix A.3. Additionally, this figure provides more context around the technical implementation of the procedure, which will be further discussed in the design section of this report.

2.1.5 Consent wizard

We faced a big challenge in how to integrate all these characteristics into our platform, without overwhelming the user. Next to that, it was important that researchers also had the ability to customize the consent procedure for their own survey, at least to a certain degree. Still, some part of the procedure would be the same in each survey, as providing (or refusing) primary, secondary and tertiary consent would be required for our system to work. This lead us to the decision of creating a wizard that takes the user step-by-step through the entire decision process. By breaking up the different steps, the user is less overwhelmed while the granularity of DIC is highlighted. The first step of this wizard is an information page, so citizens are able to make informed decisions. Furthermore, researchers can optionally add additional information about the consent regarding the specific survey. After making their decisions, the last step of the wizard shows a summary of all choices made, once again emphasizing the transparency in the process. The same wizard is also used when users want to edit their consent, as to avoid confusion.

2.2 Designing and building surveys

One of the main qualms our clients had with the current solution they use to conduct studies is the lack of flexibility. Every time they wish to create a new survey or modify an existing one, they have to get in touch with the developers of the platform. While programming every survey individually provides some benefit such as unlimited customisation, it also has many drawbacks. Individual programming means a significantly larger amount of effort needs to be put into creating each survey and a researcher would need to have a far greater familiarity with the system to be able to create a survey in this manner. Our clients do not require this level of customizability, instead valuing the flexibility to create and adjust surveys on short notice, provided the survey does meet the standards they set for it. With the current system it can take over two weeks due to the sprint duration of the development team.

2.2.1 Question builder and types

In order to solve the aforementioned issues we decided to employ a design similar to that of Google Forms. Specifically, we included a question builder as a part of the survey creation process. A researcher can create an unlimited number of questions, mark them as one-time or required, and use a WYSIWYG editor to virtually provide limitless ways of customizing each question. A researcher can pick from a variety of different question types which allow for some validation of the answer and more intuitive ways to implement certain questions, a slider type question for example being more useful than a simple number question in a place where a citizen has to give an answer on a scale from 1 to 10. In particular, we allow the researchers to pick from different question types such as *text*, *number*, *checkbox*, *dropdown*, *radio*, *slider*, and *timestamps*.

2.2.2 Survey-specific roles

In order to facilitate our strict privacy mechanisms, each survey is only accessible to researchers that were explicitly marked as participating in the research. At creation time, the creator can select who is going to be able to view or edit the survey. At the behest of the clients, we allow the owner of a particular study to mark certain researchers for read-only access to the study. This would help facilitate temporary members of the research team such as interns.

2.2.3 Submission intervals

Submission intervals are used to specify how often and at what time a survey becomes available to be filled in by a citizen. A submission interval includes several variables which are days in the month, days of the week, and times of day. Researchers may for example set a survey to be filled in twice per day, on Wednesday and Saturday. This is a necessary tool for researchers to decide on what interval they would like the survey to be filled out. The researcher specifies the interval at survey creation time, setting the variables mentioned above. The participants will also be reminded that they are able to make a submission if they indicate a reminder method during the intake process or at any point in their settings. This provides sufficient flexibility for the researchers to specify when a survey should be filled out by the participants.

2.2.4 Home page and consent

The survey builder allows the clients to design a custom home page for the survey together with study-specific consent information. This allows researchers to specify information regarding the topic of the study or consent-specific information, which citizens can view before enrolling in the study.

2.2.5 Publication and sharing

In order to facilitate the creation of short or easy-to-remember links to join a study, the researchers are free to create a custom link that is easy to share and forward by the researchers and between citizens. This means that citizens can easily remember and forward the link to others and reduces the number of errors that can occur during transcription.

2.3 Conducting surveys

Conducting surveys is a core feature of the platform. This is the flow where the data is collected and the part of the application citizens interact with the most. Because of this, a lot of time was dedicated to polishing small details in the interface. Important parts of the design include the flow for citizen signing up, the process of filling out surveys and the mechanism responsible for reminding citizens of pending surveys.

2.3.1 Joining surveys

Based our conversations with the client, there are two ways in which a citizen can join a survey. Either through a direct link provided by the researchers, or by browsing the platform for surveys they might like to join. The discovery page is the on-platform location for citizens to find new surveys. It is important that the surveys are presented in a pleasant and clear manner, and that their title and a short description is displayed so that the citizen does not need to click around much to get a basic idea of what the survey is about.

Once a survey is selected, the citizen is then guided through the consent wizard, which has been described earlier in the consent wizard section 2.1.5.

2.3.2 Filling out and viewing surveys

Filling out a survey
Filling out a survey is one of the core functionalities of our platform, and therefore it is very important that it is designed well. Since only citizens will be interacting with this part of the platform, special considerations need to be made to ensure accessibility. We took into account that scrolling is something to be avoided for citizens based on our preliminary research and what our clients told us. This means that instead of having a long survey one must scroll through, it is preferable to only have a single question on screen at a time and to have buttons to navigate from question to question. Text should be displayed in adequate font size for both the questions and the answer boxes. Additionally navigation tool can help avoid repeated clicking when a survey has a particularly high amount of questions. The other considerations we made are still with user friendliness in mind but are less specific to our user base. Such considerations include:

- A progress bar to view how far along a survey you are
- A final submit page with a confirm option so that you do not accidentally hand in the survey early
- A way to view which questions have been answered and which are still left open

Overview page
The overview page should contain all the information a citizen would want to view or edit regarding a specific survey and it should display it in a clear readable manner. There are several pieces of information which are required for the overview page based on our contact with the clients. The overview must contain general information on the study, it must contain a way to graphically view one's date and it must contain a citizen's consent information and some way to edit it. For this, the same considerations need to be made for citizens as for filling out a survey. Scrolling should be avoided, instead having several pages or tabs, all information should be clear and readable and navigation should be kept simple. We decided as well that a history tab where citizens could view specific entries they had filled in would be beneficial and our clients agreed that this would be a nice feature to have and as such it was added to the overview page as well.

2.3.3 Reminders

The clients stated that based off their experience with the previous platform, citizens appreciate being reminded by email when a survey they participate in is ready to be filled out again. For the citizens who want to be a consistent part of a study, it helps them to be get such a heads up. As such, we decided that it would be a valuable addition for our platform to have as well. This reminder can be sent via SMS and/or Email. The citizen can choose in which way they receive the reminder in their account settings. This reminder contains a personalized link to the survey page which allows the citizen to fill out the survey without logging in. If the citizen chooses to access the application through this method, the citizen is restricted to only filling out that specific survey, meaning they can not alter settings or view data. This ensures that filling out the survey is an easy process for the citizen while keeping the application secure.

2.4 Data management

After data enters the system, it needs to be managed properly. Data needs to be safeguarded from exposure to actors which do not have access to it, and citizens need to be able to view, update and delete their data at any time. Researchers need to be able to extract data as needed from the platform to process for research, while the platform should ensure data exports respect citizen consent.

2.4.1 Self-service data management

Citizens can autonomously manage any data they provided to the platform. On the survey page of any survey, citizens can view submitted entries. They may choose to delete their data or, as described earlier, to revoke their consent. All this is possible within an automated process, so that requests from citizens can be handled without further manual intervention. This helps to respect citizens' control over their data.

2.4.2 Data exports

It would be impossible to produce scientific research if there would not be a way to extract collected data in a sensible format. Data processing in the application is out of scope of the project as the researches prefer to use industry-standard tools that intake CSV files. Therefore, in accordance with the wishes of our clients, data is available to be exported in a CSV format. However, the platform places certain constrains on the exporting of data in order to prevent misuse of data by the researchers and to protect the anonymity of the citizens. The following requirements were adopted to strike a balance between citizen privacy and a researchers' freedom to perform research:

- **Scopes of data extraction**
  Researchers have to specify the scope of the data extraction. In general, there are three options for this. Firstly, after a study ends, data may be used: this is a final data extraction. Secondly, the need may arise to extract data on a preliminary basis, for example to validate the health of the data: this is a preliminary data extraction. Thirdly, a new study may be conducted that can usefully incorporate already collected data. This type of data extraction is further explained below.
- **Providing a reason for data extraction**
  For every data extraction, the researchers are required to provide a reason. This reason together with a timestamp of the data extraction is presented in the interface of affected citizens, allowing for complete transparency into the flow of data.

Exporting for other research
In the third mentioned scope of data extraction, a situation arises where data would be required for other research and therefore would need to be exported, but primary consent has not been given for the use of this data in this other research. In that case, researchers can create an extraction request with a deadline. All citizens who provided secondary consent to the survey will be notified and asked for primary consent for this new new research. As a reminder, primary consent relates to a specific research states whether a

citizen consents to their data being used for that research. Secondary consent relates to a specific survey and states whether a citizen consents to being contacted about future research that may want to use their data in this survey. After the deadline passes, the researcher can then export the requested data. Obviously, only the data of citizens that did provide primary consent will be included in the dataset. Researchers are automatically incentivized to pick a reasonable deadline: set a deadline that is too tight, and only few citizens will have had time to consider the request and give consent, while a deadline too far in the future will cause a longer delay in the extraction of data. Overall, this design decision is intended to allow the researchers to use valuable data for other studies with the consent of the citizen while keeping them informed and automating the whole process.

2.5 Communication

Another major component of our platform is the communication between citizens and researchers. As said before, transparency is something our client values very highly. Another way this is ingrained in the platform is by making it effortless for citizens to come into contact with researchers, so any unclarities are quickly resolved. On the researcher-side, it must be possible to send out announcements - usually longer messages with information about a study or Citizenlab itself - to citizens according to certain predicates. During all of this, citizens are to appear anonymous to the researchers. Lastly, the channel by which is communicated must be such that both citizens and researchers can quickly familiarize themselves with it.

2.5.1 Constraints

The main limiting factor in designing the communication flow was the lack of knowledge about the participants in their surveys from the researchers. Except for an ID, researchers have no information about the citizens, and thus no way to identify them. Still, there needed to be a way to target citizens when sending announcements. After all, sending every one of them out to all citizens in the system would not be very feasible. After discussing it with the client, it was decided that there would be a few filters that a researcher could set when sending out an announcement. For example, only targeting citizens from a specific survey that have agreed to be contacted about research outside of the scope of the current study. By only evaluating the filters in the back-end of the system, the researchers themselves have no information on to whom the announcement was sent out.

Another constraint to do with the communication flow is the dissimilarity between the wishes from the citizens and the researchers in regards to the communication channel. Citizens prefer to use email when communicating, as most of them are already accustomed to the practice and they are less keen to figure out new ways of communication.[5] On the other hand, researchers prefer to have a separate communication channel. One reason for this is because researchers will generally be receiving more messages than the average citizen. Having a separate inbox for messages from citizens will prevent their work-email inbox from cluttering up. As the client reported that communication is a very important aspect of the platform, we did not want compromise.

2.5.2 The solution

Therefore, we decided to solve the discrepancy by allowing both possibilities. We built a communication channel embedded in the platform, so it became possible for both citizens and researchers to communicate via the channel. Next to that, any messages from this channel, including announcements, are sent as an email to the involved citizens. This gives them the possibility to reply with an email, which will be collected and then displayed as a 'normal' message on the embedded communication channel. In this way, both requirements are met and communication is made as accessible as possible. Even when sending plain emails, the citizen remains anonymous for the researcher. However, we can not prevent the citizen breaking their anonymity by

[5] This claim is supported by information provided to us by the client.

providing their identity themselves. A comprehensive overview of all ways communication can be initiated is given in appendix A.4.

3 Implementation

3.1 Architecture and tooling

Careful consideration of the architecture of a software project ensures efficient development of the product. Based on the initial rough requirements, we were able to envision an architecture for the platform. First, this section will cover the architecture of the system. Secondly, our choices regarding tools will be explained.

3.1.1 Architecture

A software system consists of many lines of code, scripts, configuration files and tools; the latter of them often being referred to as the *technology stack*. These bits of software, however, are meaningless if not combined together in a proper fashion and the communication between components is not streamlined. For this reason, as part of building or implementing a software system, an architecture must be envisioned first. For the architecture of our project, we used three principles:

- **The architecture originates from requirements and design**
  Both the design of the system (such as a data model) as well as requirements imply large parts of the architecture of the system. Most requirements are at least partially satisfied by use of a specific technology. Therefore, basing the architecture on these usages makes it easier to define.
- **The architecture serves developers**
  With just a few weeks spent developing, especially within the context of this project it is vital that the architecture is ergonomic and supports development instead of hindering it.
- **The architecture follows industry standards**
  By making sure industry standards and best practices are followed throughout the project, chances of issues arising out of non-compatibility are minimized. Furthermore, by keeping close to standards it makes the project as a whole more compatible which helps when it needs to be deployed, extended or maintained.

Based on the principles above, we concluded the structure of the underlying software behind the platform must be well-defined, reproducible and according to best practices. A schematic overview of our architecture is included in appendix A.2. A short explanation on this architecture can be found below, explaining how it satisfies aforementioned principles.

Well-defined
A documented and communicatable architecture is an important feature of any software system. Most of the architecture of the platform is provided by configuration files, such as the Docker Compose YAML file. Other parts stem from the segregation of code: backend, frontend and unrelated scripts all have their own directories. Together with the overview presented in appendix A.2, it should be clear to all users of the codebase how code is structured and should be combined into a final product. Furthermore, a README file as well as many helper scripts provide enough context about how different tools can be used to carry out different actions such as generating localization files, making database migrations or maintaining the API client.

Reproducible
With just a few weeks to work with and three different operating systems among team members, no time can be wasted solving issues that arise due to the difference in environments. A good remedy for this is containerization. By creating configuration files for containerization tools, in our case a *Dockerfile*, everyone can work in the exact same environment. If a developer decides to add a dependency, package or changes the configuration, other developers can simply rebuild on their end to continue working with updates applied.

An intialization script is available, and by enforcing that script to work throughout the course of development, there always was a good basis to fallback onto should a configuration change prove problematic. Default fixtures for the database were present so any developer has a solid base to work from when it comes to initial contents of the database. Maintaining a tight containerization policy along with fixtures and initialization scripts takes a bit of extra effort, yet has the potential to save much time in the long run.

Collaboration happened through means of a Git repository. Here too, the principle of reproducibility was applied. Use was made of pre-commit hooks in Git to guarantee that every commit would satisfy certain requirements. Furthermore, the tooling used in the pre-commit hooks was able to rewrite code to a single agreed-upon style to ensure a consistent codebase. These technologies are often hard to introduce when it comes to existing projects, yet are easy to use when starting a new project, making this platform the ideal candidate for it.

Following best practices
By adhering to best practices, compatibility with other systems, operating systems, deploy environments and potential new tools to be used is promoted. Obviously, the scope of the project excluded building a solution that is fully ready to deployed. Hence, within our development environments shortcuts are taken. One example of such a shortcut is that the main container runs multiple processes, something that is usually discouraged. [2] During the development process, we made sure to write reusable and testable code. For example, in the back end we isolated business logic into small, reusable services which can easily be unit tested. In the front end, many views can be reused for both the researcher and the citizen with few changes. An example of component reuse is the survey creation wizard and the survey editing view. Both have the same fields and components, however one is linear (each "section" has to be completed before moving on). Both use the same component with different properties set, enabling or disabling enforcement of linear progress. In general, the code base complies with many best practices, such as separation of concerns and modular design.

3.1.2 Tooling

Selecting the technology and tools was one of the major decisions we had to make in the early stages of the project. We had to take multiple things into consideration, including the suitability of the technology for the task, the familiarity every team member had with the technology and the possible learning curve. In the end, the tooling we selected can be categorized in the following way:

- Containerization: Docker [4]
- Front-end language: TypeScript [21]
- Front-end frameworks: React [16] & MUI (component framework for Material UI) [11]
- Back-end language: Python [15]
- Back-end frameworks: FastAPI [5] (to serve the API) and SQLAlchemy [20] (to interact with the database)
- Job runner: Celery [1]
- Databases: PostgreSQL [14] & Redis [17]
- Authentication/Authorization Protocol: OAuth [12]
- Git [3] for collaboration, with pre-commit hooks and linters such as Black [7] (Python) and Rome [18] (TypeScript)

The most important decisions regarded the programming languages, the back-end framework and the front-end framework. Most of the team members had experience with Python, therefore using it for the back end was an easy decision. Thanks to this, we were able to start meaningful development sooner and implement features quicker than if we were to use a programming language we were not familiar with. However, most of the team had little to no experience working with a full-fledged front-end framework. Therefore, we decided to select a powerful, well-documented framework with a small learning curve. Taking this into consideration, we went with React. While React is a JavaScript framework, we also elected to use TypeScript to allow us to use types. This helps eliminate an entire class of bugs present in JavaScript. Notably, a well-supported component framework was also important with many choices. We simply went with the most popular option - MUI. The documentation was extensive with many code examples and conformed to a design standard developed by Google. Due to the popularity of MUI and the fact that material design has the backing of a large conglomerate that uses it in the design of its websites, employing it during the development of our platform would provide consistency with other websites due to, for example, similar behaviour of components. This makes the website easier to navigate and use for first-time users.

As mentioned above, most of the team was familiar with Python before the start of the project. Due to our decision to use React and make a single-page application (SPA), the back end was limited to being a simple REST API. We thought that using a full-fledged back-end framework such as Django would be overkill and unnecessarily complicate the back-end, which is why we went with FastAPI. It is simple to learn, shares many similarities with Django which many of us are familiar with but is much more lean. To interact with the database, we went with the library suggested by the Fastapi documentation - SQLAlchemy. Almost any well-documented library that protects against SQL injection attacks would have sufficed here. The same extends to Celery - we selected it as it was the most popular way to schedule and execute tasks asynchronously.

In order to streamline development of the application, we decided to containerize our application and all of the dependencies. This was especially important as our team had Linux, Windows and macOS users which would make the setup of a development environment difficult. Furthermore, developing in a containerized setup made it easy to deploy.

While the selection of other tools was significant, there were many options available with each option being equally good at the task at hand. We went with PostgreSQL as it is open source and the industry standard for relational databases. We had no need for a non-relational database and in fact benefited from the constraints of a relational database as it made development more straight-forward and less error-prone. The setup, configuration and use was also much simpler. We used Redis for message brokering for Celery and for session storage along with brute force protection. In other words, Redis was useful for non-persistent data storage. Lastly, we decided to comply with OAuth in order to allow us to integrate with single sign-on (SSO) providers used by the citizens (e.g. Google, Microsoft and Facebook).

Overall, we believe that the tooling choices we have made formed a solid technological basis for the project.

3.2 API Client

Using FastAPI, we were able to auto generate an API client for the fronted. [13]. This reduced the possibility for errors due to automatically generated types used for validating incoming and outgoing messages. It also allowed us to develop faster - a change in the back end only had to be followed by one command to change the front end.

3.3 Data modelling

Besides the right set of tools and a sound architecture, creating a well-thought out data model was absolutely vital to realize a good foundation for our project. A careless design of the data model can both cause a number of inconveniences as well as actually limit the possibilities to work with data the way we want to. As discussed in the tooling, we went with an industry standard relational database model as the need for a non-relational database did not really arise from the elicited requirements. The final data model can be found in appendix A.1. Some design choices are listed in the next paragraphs.

Users

An important part of the data model concerns the storage of users, as these are very central objects within the application. The choice was made for a single user model instead of a user model per role, since there are few configuration options surrounding permissions and a researcher, citizen or superuser therefore have an almost identical data model. Flexibility was included in terms of authentication: both two-factor authentication as well as potential integration with SSO was accounted for. SSO was never implemented, but would be highly desirable in an enterprise environment and therefore important in the data model.

Events
The event model allows for a dynamic way of storing events along with context. This way, it becomes easier to log any event within the system such as a login, enrollment or important changes to data. While it certainly would have been possible to normalize this model into many different specialized models, it felt important to strike a balance between the simplicity of the scheme and different ways in which the data can be queried. As an example, this very simple setup allowed us to quickly move on with different question types. By not normalizing this data, it becomes difficult to query on these extra parameters. This is, however, very likely not necessary as queries that target - for example - slider questions that have a maximum level of 5 are extremely uncommon.

3.4 Scalability

The front-end is completely separated from the back-end. This means the front end can be distributed on servers worldwide. This makes loading the front-end which is a relatively big task way quicker. Sending and receiving data from the back-end is a relatively small task so one central server hosting the back-end would suffice. This makes our application scalable to a large user base.

3.5 User interface

In this section we go into more detail about the implementation of the UI. Images are included to allow for a better feel of the application.

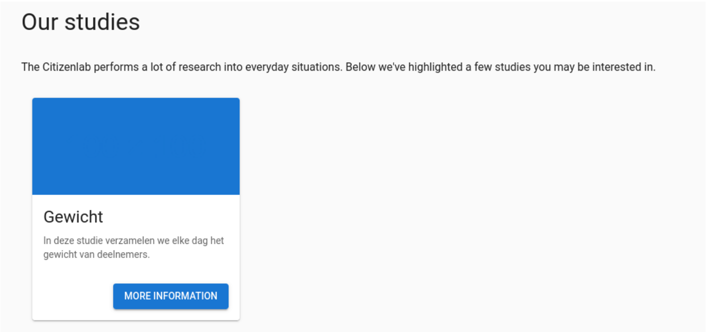
3.5.1 User-friendly design

The importance of implementing a user-friendly design was firmly established in our project proposal. We conducted literature research into the topic, so we had some foundation to base our decisions off. Next to that, the wishes and experiences of the client where obviously also taken into account.

An example of a user-friendly principle we stuck to was that we limit the text on a page to only the necessary parts. Buttons have short but descriptive names and text in the UI is provided only where explanation is required for more complicated topics, such as DIC. We do this because it has been found that elderly users commonly read every piece of text before continuing to the next page [10]. Limiting the text to only necessary parts saves time and makes it more accessible since a citizen is less likely to be confused when there is no unnecessary information to filter out. We also accompanied icons with text or use text instead of icons. This is because non-tech-savvy users often do not know the definition of most icons [19]. Another example is that we made it especially clear which text is clickable or a button, owning to the fact that people might find it hard to distinguish buttons from text. To do this, we are consistent with the button design and clearly distinguish it from normal text. On top of that, We made buttons on the citizen side quite large so they are easier to click [9].

3.5.2 Components

ReactJS allows for the creation, combining and importing of components. A component is an independent and reusable bit of code that results in HTML. Once created, they can be used as building blocks for a web page. For example, on our health platform, there are multiple pages which use the PaginatedTable



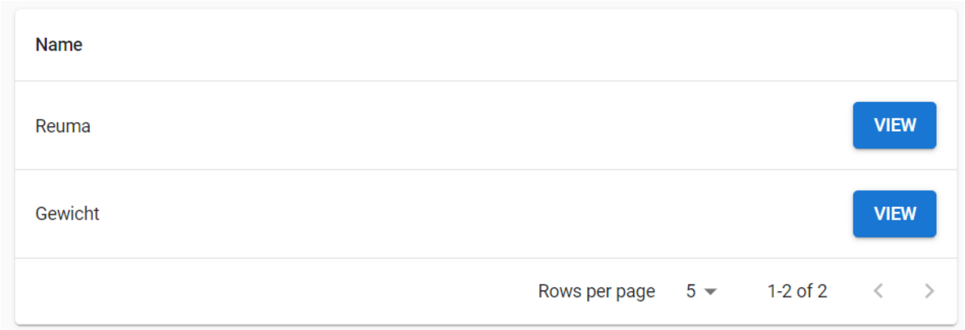
Our studies

The Citizenlab performs a lot of research into everyday situations. Below we've highlighted a few studies you may be interested in.

Gewicht

In deze studie verzamelen we elke dag het gewicht van deelnemers.

Figure 1: The discovery page showing a large ”more information” button



Name	
Reuma	VIEW
Gewicht	VIEW

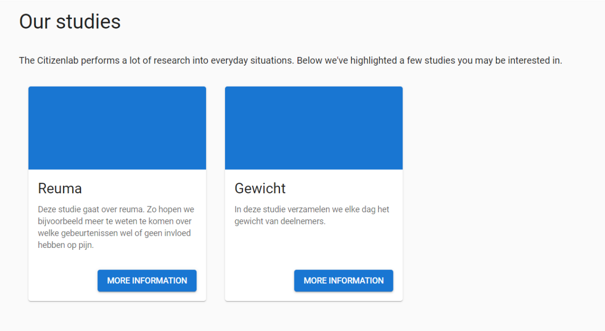
Figure 2: The paginated table component as viewed on the ”My Surveys” researcher page

component we created, which allows the user to see elements in a table with only a certain amount of entries at a time.

Components have two main benefits. Firstly, they help give the platform a consistent feel as they can be reused across different pages. Secondly, they help make the project more maintainable since a component would only need to be altered in a single place to be changed on all web pages which use it. Components can also be used to build other, larger or more complicated components. When it comes to importing them, we made extensive use of the MUI library which contains many useful UI components. This library follows Google Material Design guidelines [8]. These guidelines specify the looks and feel of objects and text on a web page. Since these guidelines are followed widely among many different websites, it means that users likely have some sense of familiarity with the design of these components. Additionally, these components come with many options for customizing them to the specific needs we had for our platform. Using MUI as a basis, we were able to build our platform with a consistent and clean look, while greatly aiding its maintainability.

3.5.3 Joining a survey

A citizen is able to navigate to the consent page to browse surveys to join. Alternatively, they can join a survey directly through a link. On the discovery page, each survey is displayed in a box with a description and an optional image. A citizen is able to select any survey they wish to join by pressing the ”more information” button and proceeding to the consent wizard which is described in the next section.



Our studies

The Citizenlab performs a lot of research into everyday situations. Below we've highlighted a few studies you may be interested in.

Reuma

Deze studie gaat over reuma. Zo hopen we bijvoorbeeld meer te weten te komen over welke gebeurtenissen wel of geen invloed hebben op pijn.

Gewicht

In deze studie verzamelen we elke dag het gewicht van deelnemers.

Figure 3: The discovery tab view of a citizen

3.5.4 Consent wizard

A citizen can give primary, secondary and tertiary consent by going through the consent wizard. They encounter it whenever enrolling in a study, as well as when editing their current consent. In the following paragraphs we will explain why the wizard has a user-friendly design and how it guides the citizens towards making their own decision. For reference, the different steps discussed are also shown in Figure 4.

Information
Because the basis for informed consent is being informed, step one of the wizard is dedicated entirely to information regarding the consent. On this page, the researchers can add bullet points which describe what the consent for this specific survey exactly entails. To counter an overabundance of information on the screen at the same time, the description of the point will only be shown if the citizen interacts with it. The fact that a citizen can press on these point is indicated. After having read all information, the citizen can continue to the next phase.

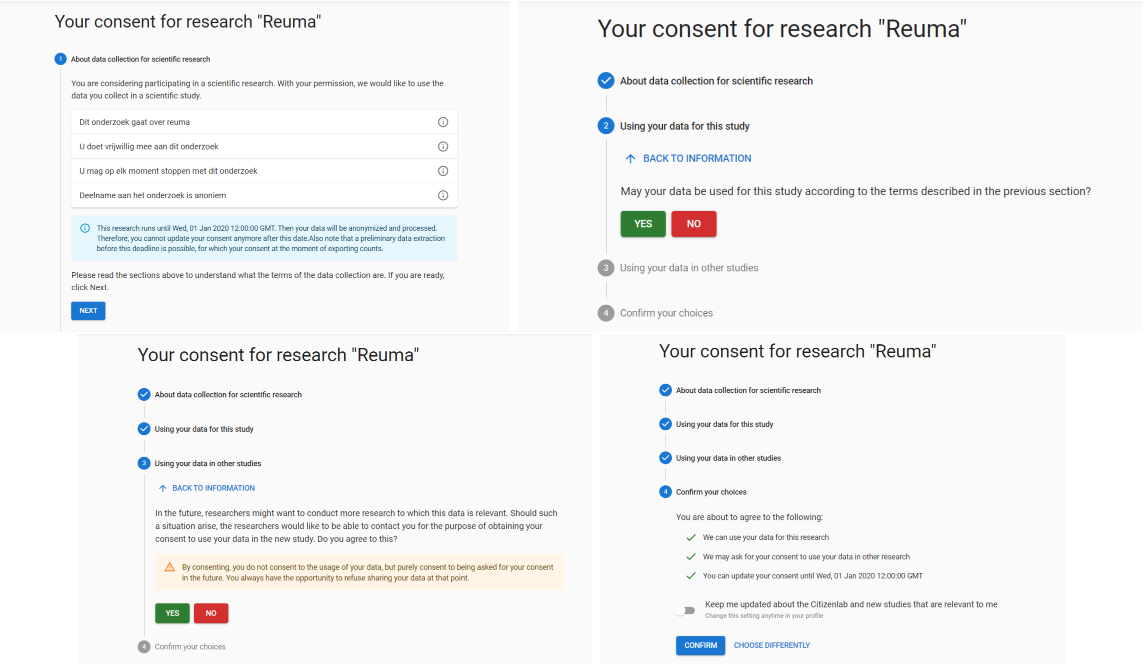
Giving consent
As explained in Section 2.1, there are many different types of consent. The first question allows a citizen to specify whether they want to give primary consent. Obviously, it is also explained what primary consent entails. If the citizens answers yes, they continue to the question regarding secondary consent. Since providing only secondary consent is illogical, the secondary consent question is skipped in case no is answered. In the fourth and final step, a summary of all given answers is given. Next to that, citizens can also indicate whether they want to give tertiary consent here. When satisfied with their answers, citizens can then save their consent settings by clicking the confirm button. At all times an option is provided to go a step back and change the consent.

3.5.5 Filling out surveys

Some implementation details surrounding the UI for filling out a survey are discussed. The components discussed are shown in Figure 5.

Navigation
While filling out the survey, each question is shown on a different page on the screen. There are two ways of navigating between questions that are available to the citizen. There is a go-to question button and there is a progress bar with buttons. The go-to question button exists so that a citizen does not have to click back many times to edit an answer to a previous question. When trying to skip over a required question, a warning is displayed, explaining why this action is not allowed.

- **Go-to question button**
  When this button is pressed, a pop-up appears with buttons that direct to each question. The color of the button indicates if the citizen already answered the question or not. The buttons also display



Your consent for research "Reuma"

1 About data collection for scientific research

You are considering participating in a scientific research. With your permission, we would like to use the data you collect in a scientific study.

- Dit onderzoek gaat over reuma ⓘ
- U doet vrijwillig mee aan dit onderzoek ⓘ
- U mag op elk moment stoppen met dit onderzoek ⓘ
- Deelname aan het onderzoek is anoniem ⓘ

ⓘ This research runs until Wed, 01 Jan 2020 12:00:00 GMT. Then your data will be anonymized and processed. Therefore, you cannot update your consent anymore after this date.Also note that a preliminary data extraction before this deadline is possible, for which your consent at the moment of exporting counts.

Please read the sections above to understand what the terms of the data collection are. If you are ready, click Next.

NEXT

Your consent for research "Reuma"

✓ About data collection for scientific research

2 Using your data for this study

↑ BACK TO INFORMATION

May your data be used for this study according to the terms described in the previous section?

YES NO

3 Using your data in other studies

4 Confirm your choices

Your consent for research "Reuma"

✓ About data collection for scientific research

✓ Using your data for this study

3 Using your data in other studies

↑ BACK TO INFORMATION

In the future, researchers might want to conduct more research to which this data is relevant. Should such a situation arise, the researchers would like to be able to contact you for the purpose of obtaining your consent to use your data in the new study. Do you agree to this?

⚠ By consenting, you do not consent to the usage of your data, but purely consent to being asked for your consent in the future. You always have the opportunity to refuse sharing your data at that point.

YES NO

4 Confirm your choices

Your consent for research "Reuma"

✓ About data collection for scientific research

✓ Using your data for this study

✓ Using your data in other studies

4 Confirm your choices

You are about to agree to the following:

✓ We can use your data for this research

✓ We may ask for your consent to use your data in other research

✓ You can update your consent until Wed, 01 Jan 2020 12:00:00 GMT

○ Keep me updated about the Citizenlab and new studies that are relevant to me
Change this setting anytime in your profile

CONFIRM CHOOSE DIFFERENTLY

Figure 4: All steps of the consent wizard

which questions are required to be filled out by showing a little star next to the number. This way, the user has an easy way of navigating and an overview of the survey.

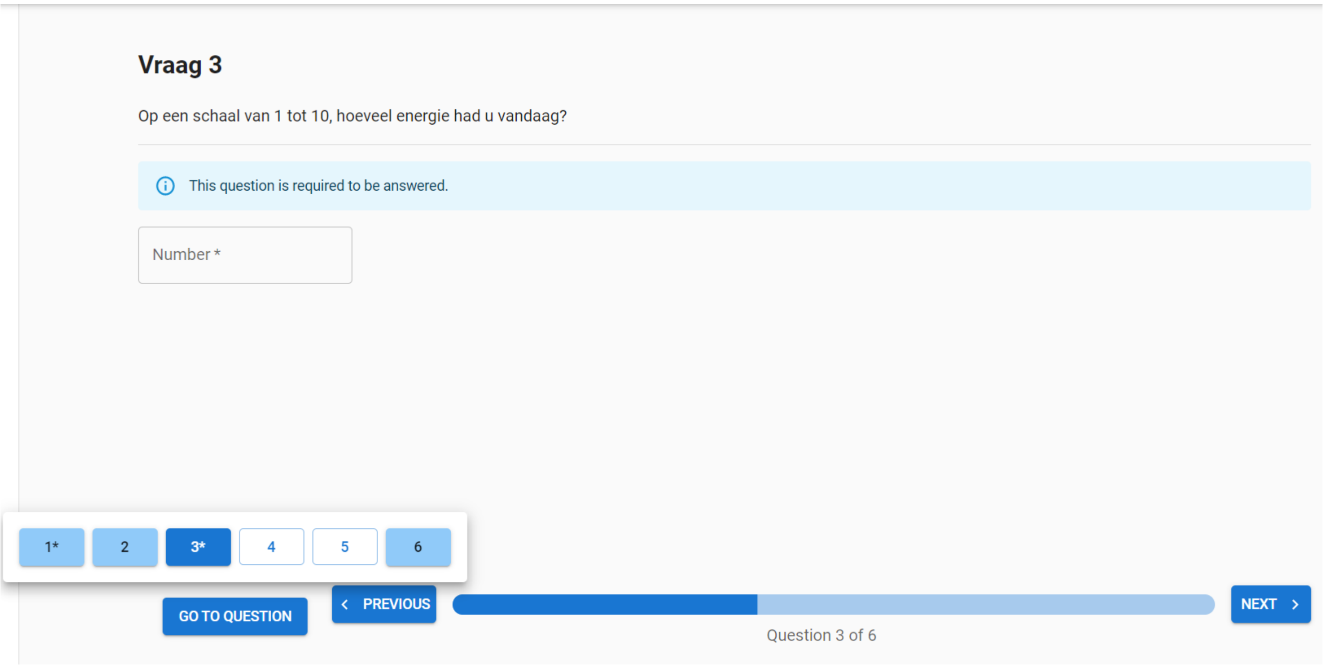
- **Progress bar with buttons**
  The progress bar is always visible at the bottom of the screen and displays how far the user is with filling in the survey. Below this progress bar another indicated is placed which tells you on which question out of the total amount of questions you are. On the left and right of the progress bar, there are buttons to easily navigate to the next or previous question.

Required questions
As already mentioned, a survey can contain required questions. To keep the flow of filling in a survey simple, a citizen can not skip a required question and later come back to that question. When a question is required, this is clearly shown in a small information banner above the input field. When the citizen tries to continue without filling in the required question, a clear error message is shown. The required banner and the error message clearly indicate what is happening at all times to prevent confusion and frustration for the citizen.

3.5.6 Communication dashboard

Like previously mentioned, the communication dashboard provides both citizens and researchers with an easy way to communicate while ensuring that the anonymity of citizens remains intact. Citizens are also able to communicate via email. As researchers would only be able to use the communication dashboard to communicate, it was vital that proper care was taken to ensure a user-friendly experience. From the client we gathered that they would likely also want to use the dashboard from their phone, making a responsive design a necessity. A selection of images showcasing the communication dashboard can be found in Figure 6.

Features
To allow for more accessibility and functionality, some additional features are added on top of the base functionality. These are the following:



Vraag 3

Op een schaal van 1 tot 10, hoeveel energie had u vandaag?

This question is required to be answered.

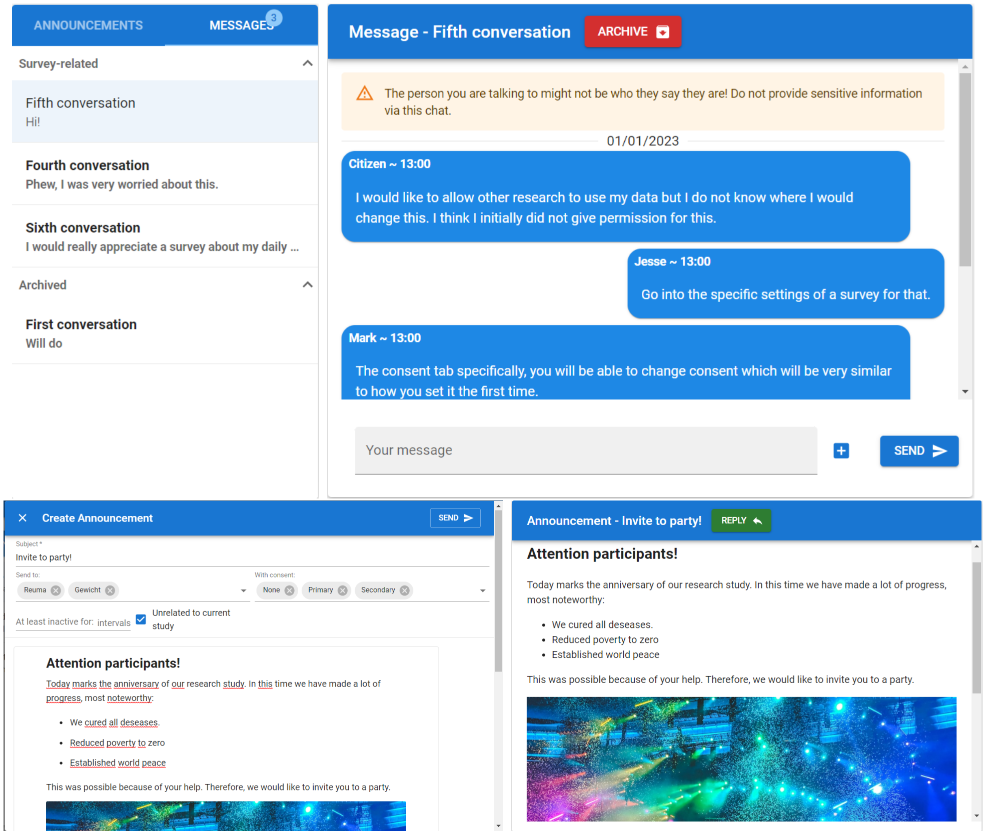
Number *

Question 3 of 6

Figure 5: Filling out a survey

- **Unread messages highlight**
  Both citizens and researchers can effortlessly see which announcements and conversations have not been read yet by simple highlighting of said messages in the overview tab. When viewing the message, the highlight will disappear. Next to that, when there is one or multiple announcements or conversations that have not been read yet, a notification badge showing the number of unread messages will be displayed at the top of the overview tab as well.
- **Ordering based on last message sent**
  Conversations are ordered based on the timestamp of the last message in the thread.
- **Categorizing conversations**
  Researchers automatically have their conversations categorized as either a general, survey related or announcement related conversation. This makes it easier to look for a specific conversation. It is also possible to simply collapse an entire category, making the overview tab less cluttered.
- **Archiving conversations**
  Researchers are able to archive conversations, which will remove them from the standard conversation categories. Instead they will only appear under a separate archived section. This gives researchers the ability to keep their inbox organized. If a citizen sends another message in an archived conversation, the conversation is opened again. Researchers can also manually reopen conversations.
- **Rich formatting of message**
  While the standard input for a message is a simple text box, users are also given the ability to open a WYSIWYG-editor, allowing for much richer formatting.
- **Anonymous**
  Regardless of the method of communication the citizen chooses; they remain anonymous to researchers at all times. This way, researchers can assist citizens in using the platform while preserving privacy and scientific integrity.

- **Responsiveness**
  A responsive design ensures that the dashboard is usable on screens of all sizes. For example, when using a phone, buttons lose their text to compensate for lack of room and both the message select-panel and content display take up the entire screen.



Fifth conversation
Hi!

Fourth conversation
Phew, I was very worried about this.

Sixth conversation
I would really appreciate a survey about my daily ...

First conversation
Will do

The person you are talking to might not be who they say they are! Do not provide sensitive information via this chat.

01/01/2023

Citizen ~ 13:00
I would like to allow other research to use my data but I do not know where I would change this. I think I initially did not give permission for this.

Jesse ~ 13:00
Go into the specific settings of a survey for that.

Mark ~ 13:00
The consent tab specifically, you will be able to change consent which will be very similar to how you set it the first time.

Your message

Subject *
Invite to party!

Send to: Reuma, Gewicht

With consent: None, Primary, Secondary

At least inactive for: intervals

☑ Unrelated to current study

Attention participants!

Today marks the anniversary of our research study. In this time we have made a lot of progress, most noteworthy:

- We cured all deseases.
- Reduced poverty to zero
- Established world peace

This was possible because of your help. Therefore, we would like to invite you to a party.

Attention participants!

Today marks the anniversary of our research study. In this time we have made a lot of progress, most noteworthy:

- We cured all deseases.
- Reduced poverty to zero
- Established world peace

This was possible because of your help. Therefore, we would like to invite you to a party.

Figure 6: A selection of screenshots from the communication dashboard

3.6 Deployment

Developing a completely deployable system has never been within the scope of the design project. As explained in previous sections, however, we do believe a solid technical foundation has been estreport/figs/announcement-creation.pngablished. Therefore, it is useful to make some remarks regarding required changes in order to run the application reliably in a production environment.

3.6.1 Security

A logical starting point is assessing the security of the application. Since security should not be tied to any particular environment, it has been accounted for from the start and embedded within the application code. This mainly amounts to using best practices, but a correct separation of concerns also helps in this regard. Apart from ensuring the target infrastructure is secure in itself, the only necessary changes would be to use safe values in the main configuration file. Safe defaults were chosen to facilitate this.

3.6.2 Supporting services

The platform requires a few supporting services which in the develop environment are being ran as separate containers. Examples of this include the database (PostgreSQL), Redis and more. For a full reference, see appendix A.2. These services should not be temporary containers, but rather set up in a reliable way and be accessible to the application core.

3.6.3 Core

The core of the application, consisting of a front-end and back-end, can be deployed in numerous ways. Since the front-end consists purely of static JavaScript, CSS, HTML and a few assets, this can be deployed to any CDN or hosting platform. The back-end, written in Python, has to be run on an application server that both supports the dependencies as well as is able to connect to the support services mentioned earlier in this section.

4 Process

We divided the process of designing the application into 5 subsections. Gathering the requirements, Lo-Fi prototype, Hi-Fi prototype, MVP, Final product. These subsections are listed in chronological order and each represents about 2 weeks.

4.1 Requirements engineering

With only a project description of a single page to work with, the first step in building the platform was eliciting requirements from the client. The first two weeks were mostly spent orienting, eliciting requirements, researching and brainstorming. Eventually, a project proposal arose out of the various activities and gathered knowledge.

4.1.1 Whiteboard sessions

In the first week, many brainstorming sessions were organized to gather ideas, start shaping the platform and defining the outer boundaries of what was to be built. In total, over the course of two weeks, about six meetings were held which would last half a day. Overall, the topics discussed amounted to:

- Practical issues related to the organization of the project, such as selecting a supervisor
- Developing an outline for the proposal
- Preparing the first meeting with our client
- Deciding on our process of collaboration
- Performing exploratory research into relevant topics related to the project

The meetings did not follow a structure or agenda; rather topics were addressed in the order they became relevant. Practical issues such as selecting supervisors and setting up procedures surrounding the project took precedence in the first week while in the second week brainstorming sessions that discussed the platform on a conceptual level became more important.

4.1.2 Client interview

To start the requirement elicitation process, we organized an interview with the client. This interview was set out to last about two hours. The main goal of the interview was to elicit all important general requirements, at least in a rough sense. In other words: it was supposed to be a very broad interview, but not necessarily very deep. The interview consisted of four phases.

The first and last phase of the interview were not part of the requirements elicitation; they were more about general topics such as exchanging information about both the Design Project as well as the Citizenlab. The second phase was aimed at discovering concrete *goals* of the client, with the third phase being an extension of that, which would produce actual *requirements*.

To prepare the second phase, we established a few categories of requirements we expected would be relevant for the platform. On the one hand, scoping the interview to a few set categories risks introducing a bias in the results; the client may be tempted to only discuss categories the interviewers deem relevant and not mention other important information. On the other hand, it is practically impossible to ask about such an extensive set of requirements in an unstructured manner. Therefore, we opted with a semi-structured interview where we already defined categories of questions. During the interview, goals the client mentioned were categorized. This information was subsequently used in the third phase.

After collecting collecting the goals the client wanted to achieve, it was time to extract concrete requirements from these goals. Instead of asking broad questions, more targeted and detailed questions were asked based on the inputs the client provided in earlier stages. These requirements were then used in our project proposal.

Results
Everything clients mentioned during the interview was processed into a list of requirements which is included as an appendix to the proposal.[6] Even requirements that were out of scope were included; this ensures completeness.

4.1.3 Desk research

From the beginning, it was very clear that certain topics such as accessibility would play an important role in our product. For this reason, it was decided that it would be beneficial for the project to review some literature in these areas. This was done over the course of the two weeks. The knowledge was shared among team members and documented within the project proposal.[7] This knowledge also was applied in the design: good examples of this practice can be found in section 3.5.

4.1.4 Project proposal

The end product of these two weeks consisted of the project proposal as it was presented to our clients. This proposal is included in appendix F. Shortly after finalizing it, it was approved by the clients. Throughout the project, this proposal served as the main document describing our end goals. We committed to the requirements listed in the proposal and were guided by the research and context that is included in the proposal. Furthermore, a timeline was part of the proposal which provided guidance on when to carry out the various activities.

4.2 Lo-Fi prototype

During the third week of the project we worked mainly on the Lo-Fi prototype. Our Lo-Fi prototype was created through the use of Figma. [6] Initially, we brainstormed on what all the required pages would be for the three user roles, those being citizen, researcher and superuser. Once we had a clear idea of all the required pages for our application, we created a few components that could be reused between pages, such as the navigation bar, and then proceeded to divide up all the pages of the Lo-Fi prototype between our group members. In our Lo-Fi prototype we included example pages for both a PC screen view as well as a phone screen view as our clients communicated an interest for having phone compatibility for the platform.

The final result of our Lo-Fi prototype included roughly 30 different pages which together gave an overview of what the final application would look like. Additionally, it allowed us to reason and communicate better with the clients about some of the functionalities of the platform and how they should be implemented. We

[6] Appendix F
[7] Appendix F

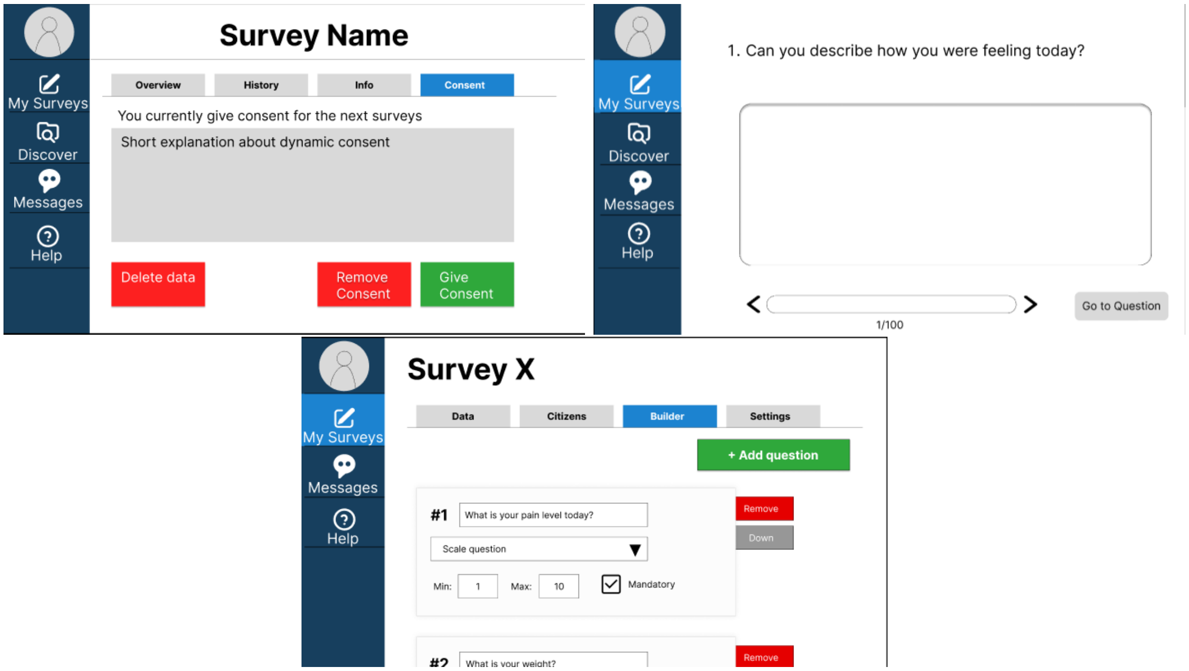
showed the Lo-Fi prototype to our clients during a meeting. The clients were mostly positive about the design and the range of features included. Some feedback was given based on our Lo-Fi prototype as well.

Notable feedback was given as follows:

- The question editor was satisfactory to the clients
- Mandatory questions is an important feature to include to avoid gaps in the data
- Citizens should not be able to edit answers after the fact
- The clients were happy with the design of the informed consent page
- Information in the informed consent page should be concise but complete, possibly allowing for drop down menu's or hover over information boxes if citizens wish to read all details
- Whether a survey can be found on the discovery page should be configurable

Notably, most of the feedback we got from the citizens was regarding certain functionalities and options instead of just design. This was mostly because overall, the clients were quite happy with the general look and feel of the website. The fact that the Lo-Fi prototype showed a clear initial picture of the platform did however help the clients with spotting some features that would need to be added for the platform to function the way they want it to. Additionally, the Lo-Fi prototype helped us to more clearly define the states a survey can be in, using the survey settings page as an example. Those states being published or not, discoverable or not, started or not and whether it can be filled out after the study has already expired.

All feedback our clients had in response to the Lo-Fi prototype was incorporated into our work during the rest of the project and the features listed as well as the UI considerations have all been implemented.



You currently give consent for the next surveys

Short explanation about dynamic consent

1. Can you describe how you were feeling today?

#1 What is your pain level today?

Scale question ▼

Min: 1 Max: 10 ☑ Mandatory

#2 What is your weight?

Figure 7: Consent overview tab, filling out a survey and survey builder of our Lo-Fi prototype

4.3 Hi-Fi prototype

After finalizing the Lo-Fi prototype, the Hi-Fi phase started. In this phase, we started building the actual application with the approved and adjusted design from the Lo-Fi prototype.

4.3.1 Data model

We started by designing the data model as this is a determining factor within the project structure. We decided to let each of the team members come up with their own design of the data model. This allowed us to easily compare any differences. Interesting discussions arose, which helped us make well-informed decisions. The activity also had other advantages. Each team member was forced to consider every aspect of the application. By doing so, a better understanding of the complete project was formed. Differences in the interpretation of specific sections of the application were promptly highlighted by the designs. These differences were ironed out and going forward, the entire team had a unified view on the objective at hand.

4.3.2 API

After having the data model in place we continued with creating stubs for all API endpoints we anticipated we would need. This was not necessary for the Hi-Fi prototype, however it was helpful in structuring the front end code. It also came in handy when we got around to implementing the back end and integrating it with the front end. As part of the API stubs, we defined the URI, the request body, the response body and query parameters. With this information, the framework we chose to use for the back end was able to auto generate documentation for the API.

4.3.3 Front-end

We started constructing the Hi-Fi prototype by rebuilding the Lo-Fi prototype in React using Typescript and MUI. This tech stack allowed us to rapidly develop an accessible and responsive interface.

We received useful feedback during a meeting with the client where we showcased the Lo-Fi prototype. During the development process of the Hi-Fi prototype, we made sure to incorporate said feedback into the design. Due to the fact that the application had to be mobile friendly and because of slight differences between the components we used in the Lo-Fi prototype and the ones provided by MUI, there were slight differences between the Lo-Fi prototype and the Hi-Fi prototype. However, most of the time the differences were minor (for example, the default colour scheme).

Due to the fact that we had not yet started on developing a functional back end, we employed mock data to emulate the back end and allow for some data to be displayed to test our views and host demonstrations for the clients.

As we have mentioned before, the Hi-Fi prototype was the final interface of the application, therefore images of the Hi-Fi prototype can be found in the 3.5 section of this report.

4.3.4 Retrospective I

The first retrospective took place in week 5 of the project. At this point, we were able to show a concrete realization of some of the core requirements on the front-end of the application. The clients expressed that they were happy with the progress so far. A few labels were deemed confusing and the requirement of a one-off question was added. We also discussed about the different roles needed in the system. The conclusion was that they needed a superuser role, a researcher role, and a citizen role. For a survey, they needed an editor and a viewer role.

4.4 MVP

It is hard to tell when a product is to be regarded as *minimally viable*, yet we had a clear vision for this step in the process. We wanted to have a bare version of the platform that was deployable, maintainable and contained the core of the requirements with good means to extend these to all target requirements.

4.4.1 Putting things together

When constructing the MVP, the Hi-Fi prototype was used as a basis. Since this was already expressed in code, it was ready to work with. By starting to integrate both the Hi-Fi prototype as well as the API stubs and building functionality, we quickly started moving towards a functioning product. It was at this time the database was modelled and containerization set up. With all team members working on the product at the same time, this presented unique challenges. However, by communicating well and using the right tools in the right way no major issues arose during this stage of the project.

4.4.2 User testing

Near the end of this phase in our project, we performed some user testing on the MVP. Before starting this, a request was made to the ethical committee of the UT which was approved. The user testing was mainly done with relatives since setting up a test with the actual target group, was deemed impossible in such a short time. A test plan was created to ensure important parts of the application would be tested as well enable consistency amongst test. We tested the core flows on the citizen-side of the system, such as: joining a survey, filling out a survey, viewing the submitted data and communicating with the researcher. The received feedback was implemented in the final phase of the project.

4.4.3 Retrospective II

In the second and last retrospective in week 7, we spent most time showing features to the client. While nothing is set in stone, all parts of the platform were at least partially implemented so room to make large changes was very limited. This being the first retrospective where this was the case, it was an interesting moment to see if requirements were being implemented correctly from the perspective of the client. The clients were once again satisfied with the process and product so far, and no major changes had to be made. Therefore, work carried on leading up to the final product.

4.5 Final product

4.5.1 Goals

The final product stage of the project was the ending stage of the design and development process. As we were due to have our final meeting and showcase of the product for the clients on April 13h, 2022, any changes made past this point would not have any use or impact in the context of the design project.

Because of the major progress we have made in the MVP stage, we had the vast majority of requirements marked as ”must” on the MoSCoW scale implemented in both the back-end and the front-end. In this stage of the project we were focused on polishing existing features and implementing lower priority tasks.

4.5.2 Challenges

While the majority of challenging and domain-specific features were implemented in the MVP, we found this stage much more difficult. While our MVP was usable, it lacked certain generic features like the ability to change your password. While details like this are minor, they greatly contribute to the usability of the platform and its ability to be deployed to be used as a functioning platform.

The scope of the project was large. Striking a balance between what is possible to implement in a few weeks and a long list of requirements has therefore been a challenge. We opted for a *broad* approach, attempting to cover a lot of requirements, but not always going into great detail. To illustrate, it would not be unrealistic to set up an entire ten-week project just to develop the communication section. By approaching the matter from a more shallow angle, we attempted to succeed in giving a very good and solid impression of possibilities as well as providing a good basis for future work. The actual challenge here to overcome was to only allow features into scope that could actually be completed to a reasonable degree. That does not imply feature completeness, but it does mean we avoided unrealistic requirements and thus prevented the product from containing too much rough and unfinished work.

4.5.3 Final showcase

The final showcase was the last meeting we had with the clients. The purpose of the meeting was to allow the client to freely use the application and evaluate the product. Due to the large number of features of the platform, we decided to give the users a number of tasks to fulfil in order to demonstrate the potential of our platform. In other words, this showcase took a format similar to a semi-structured user test. While we have gathered feedback from the clients over the course of the project, this was the ultimate litmus test of whether we have fulfilled all of the core requirements. There were some minor issues that we missed during system testing partly due to the different hardware used by the client. For example, due to a likely issue with floating point arithmetic in a library we were using, a graph rapidly shrunk upon opening a particular web page on the client's computer. However, in the end, the feedback from the clients was that we have exceeded expectations. All issues encountered during the demo have since been fixed.

5 Testing

Different phases of the project called for different ways of validating our methods. Within earlier phases of the project, we made use of human subjects to user test ideas and prototypes. During the implementation phase we used technical means to verify the system is behaving as expected.

5.1 User Testing

5.1.1 Ethics approval

To be able to carry out our user testing, we filed a request with the Ethics Board[8] early on during the project to verify that our purposes and usage of information for our user testing was ethically sound. Our request was accepted by the Ethics Board which meant we could proceed with our user testing further down the project.

5.1.2 Testing setup

The user testing itself was carried out once we were working on the MVP. In order to carry out the user tests, we created a user testing document. [9] The user testing document includes the flows a user should be taken through as well as the questions they should be asked afterwards of how they experienced that flow. For our user testing, we specifically focused on the citizen side of our health platform as for the researcher and superuser side we received much feedback already throughout the project from our clients. The user group which we tested consisted of our parents, as they are roughly in the same age range and have roughly the same level of experience using technology as the citizens of Citizenlab do.

5.1.3 Results

Each user would be guided through the different flows a citizen might go through when using the application. These being: joining a survey, filling out a survey, viewing a survey and its settings, and viewing ones own

[8] Ethics Board of the EEMCS faculty of the University of Twente: https://www.utwente.nl/en/eemcs/research/ethics/

[9] Appendix C

account settings. From our user testing we obtained several useful points of feedback which we considered for our final design.

Several notable findings from the user testing were:

- The naming of the ”discover” page caused some confusion
- Users understood the informed consent procedure - provided that they took time to read it - and were able to put what it meant into their own words
- The go to question button caused some confusion due to its positioning and the survey questions itself being too empty
- Users missed a button from the history page to the overview page

These were the results that were shared between multiple users. We were pleased to see that in terms of the information provided, the users either stated it was clear or could give a satisfactory reproduction of what the information meant in their own words.

Some smaller difficulties were experienced here and there, they were however not shared between users and therefore not considered to be as significant. Those results are as follows:

- A user found it initially unclear where to find the personal settings
- A user experienced confusion as to the exact meaning of revoking all consent and how large of an impact it would have
- A user found it unclear what the notification settings for their personal account meant

In our final design, we accounted for the feedback which was shared by multiple users by renaming some components and adding a back navigation to the page it was missing from.

5.2 Test plan

Due to the short time frame of the project, we had to carefully consider how we are going to test components of the project. While in an ideal world we would employ test driven development, it would have eaten into precious development time. Therefore, we decided to focus on system testing with unit tests only for the most sensitive components of the platform.

5.2.1 Testing and security

As we have mentioned multiple times, security is one of the most important aspects of the platform. Coincidentally, due to the fact that business logic is decoupled from other application parts and implemented within services, it is easy to unit test. Most unit tests focus on testing critical security features.

Due to the way our application is structured, one must simply define an endpoint as protected and which user types are able to access it. Upon a request being received, logic is triggered to check that the request is allowed to proceed. Therefore, in order to validate that each endpoint is secure is as simple as unit testing the shared authorization logic.

Furthermore, out system employs a brute force protection system that will lock out users and IP addresses that try to authenticate or authorize with the server too many times in quick succession. This component is also vital to the security of our platform and is also unit tested.

5.2.2 Data export

After a user authenticates, the session token they they bundle with each request associates them to the request allowing endpoints to return data only related to them. However, components such as the data export module allow researchers to export the data of every consenting citizen for research purposes. As the logic is more complex and separate from the aforementioned authorization mechanism, this logic has to be tested in order to ensure that the data of citizens participating in other studies or citizens that have not consented for their data to be used for research purposes. As this module is well-decoupled, it is also trivial to test it focusing on the correct inclusion and exclusion of citizens.

5.2.3 Other components

The majority of the rest of the application consists of the front end, which based on our repository statistics makes up around 60% of overall code. We decided that due the failure mode of these components being acceptable (in the worst-case the application will crash, display unexpected error messages or be displayed incorrectly) the front end should only be tested through system tests only.

6 Reflection

Learning from projects such as these is only an achieved goal when a proper reflection took place. In this section we reflect on the risks we encountered and our collaboration. Furthermore, we will be presenting some possibilities for future work.

6.1 Risk

With any new development, risks arise, and ours is no exception. We have built a platform for collecting data, including personal and even medical data. Privacy, security and ethical considerations play a prominent role. In our proposal, we defined a few categories of risk.[10] Besides the operational risks - which remain unchanged - we revisit planning risks and implications to privacy and security in this section. Furthermore, some attention is directed towards ethical considerations as a new category.

6.1.1 Planning risks

Fortunately, the planning risks as identified in the proposal never became reality. The project progressed as expected and encountered no major problems.

6.1.2 Privacy and security risks

Although all risks contain an ethical component, in the following paragraphs, we try to focus more on risks related to the platform-functionality.

Account takeover
A big risk to our application is an account takeover. This event can happen for many reasons. If a user is careless about leaving their password around or even tells another person their password, an attacker can use it to gain access to the account. Another risk is if a user uses the same password for multiple platforms which can be leaked in case of a data breach. Even using a very simple password can be the cause of an account takeover. The consequences of this happening differs for the type of account being taken over. If a citizen account is taken over, all data of that citizen can be accessed and even easily downloaded. Although this is a big problem, an even bigger problems occurs when a researcher account is taken over. This will allow an attacker to extract all data from the surveys the researcher is a part of. In case a superuser account is taken over, the attacker is even able to make changes to the platform that affect every user or create a new researcher account that can trick citizens who believe the researcher can be trusted.

Because of the big risk an account takeover poses, we have a number of policies in place to mitigate some of the risks. The risk of a user using a very simple password and an attacker bruteforcing the account is mitigated by the requirements we set for a password. These make it so a user cannot set a very simple

[10] Appendix F

password like '123' which is easily cracked. Another way we discourage bruteforcing is by timing out user accounts after a number of invalid authentication attempts. We do the same for IP addresses. If a number of invalid authentication attempts have been made from an IP address, the IP address is timed out and has to wait a while before being able to try again. To make authentication less reliant on a single password, we allow for TFA (Two Factor Authentication). If this is enabled, a user is unable to login with just their password, they will have to prove it is them by agreeing to a notification on their phone as well. Since researcher and superuser accounts pose a higher risk in case of an account takeover, TFA is enabled by default for these accounts and cannot be turned off. Lastly, as a precaution to an attacker being able to pose as a researcher and trying to trick a citizen into giving them personal information, we always inform the citizen that there is never a need to provide personal information via the chat.

Unauthorized survey submissions
When a citizen has joined a survey and they agree to receive reminders for it, they will get an email containing a personal link for filling in the survey whenever it is available. For convenience, we allow citizens to fill in the survey directly without logging into their account when they use this link. Of course, this poses the risk of malicious agents coming into possession of the link, allowing them to fill in the survey on the citizen's behalf without having to know their password. However, they are unable to view any of the user's past data. Because of the low impact, we decided to make this trade-off between accessibility and security.

Unauthorized joining
Very similar to the previous risk, each survey has a unique link which directs a citizen to its information page. From here, citizens are able to join the survey. Surveys can be set to private, this means that it does not show up on the discovery page and citizens can only join it using this unique invite link. If a user guesses the link or finds it by some other means, they would be able to join the survey even though a researcher might not want them to. Again, this risk is small, as it will only allow for the addition of data, which can be filtered out after the extraction. Still, to mitigate the risk of being able to guess the link, we allow researchers to define their own unique link, so they can make it as complex as they want. This can make it hard if not impossible for users to guess the link.

Administrator lockout
The superuser does not have a way to recover their account when they forget their password. Along with the obvious risk of losing data related to the account, if all super users forget their passwords Since the number of admins is meant to be kept low, as an increase in admins is an increase in risk, there is a possibility that all admins are at some point locked out of their account. This would mean a loss in functionality, as site-wide settings can no longer be updated and new researchers/admins can no longer be created. However, the maintainers of the platform will always be able to add new superusers directly to the database, making this risk more of an inconvenience than anything else. Still, a good practice would be to always have a minimum of two active superuser accounts.

Citizens acting inappropriately
Moreover, there is the risk of citizens acting in an inappropriate manner. This includes citizens filling out surveys in an incorrect manner on purpose. Citizens are aware that they are anonymous when filling out surveys which might make them more likely to intentionally provide wrong information. We leave it to the researchers to filter out data they deem untrustworthy, just like it would be had they conducted the survey in any other manner. Another problem, more specific to our platform, is the ability for citizens to spam the communication channel. There is no direct way to counter this behaviour, for example by blocking communication for a citizen. However, we do make sure that researchers do not get emails whenever they receive a message, as opposed to citizens, which mitigates the risk of their email inbox filling up. A feature researchers can use to unclutter their in-platform inbox is the ability to archive conversations. Although a conversation is unarchived when a new message in that conversation is send, it at least allows for cleaning up the inbox after the barrage of messages has stopped.

Malicious researchers
Finally, a risk we believe to fall outside of the scope of our platform is the risk of malicious researchers.

Looking back at the DIC, explained in section 2.1, researchers are expected to indicate when they plan to extract data for a study other than the primary study. Then, citizens who gave secondary consent are contacted to ask for their consent once again. However, if researchers refrain from indicating this fact in the export request, they will simply extract data from all citizens who gave primary consent, without an additional check. This is obviously unethical (and unlawful), but the problem is not restricted to our platform. Any researcher conducting their own survey is able to use the data for other purposes. A breach of confidence of this severity does not fall within the scope of the platform since researchers are expected to act in an ethical manner.

6.1.3 Ethical considerations

Within the scope of the project, there are many ethical issues to consider; are we able to to guarantee the safety of the data, and if not, does the risk outweigh the benefit? Do we even have valid reasons to justify the collection of any data? Even if these questions cannot be definitively answered, it is ever valuable to analyze them and see if we can come to some sort of consensus. A reflection on these and other ethical considerations can be found in D.

6.2 Teamwork

Going into the project, we already knew one another from previous modules, which meant we skipped past the team building phase and instead immediately started the project. During the project, we met up most days to work and communicate. This helped us maintain a clear picture of our progress as the project went on.

During these sessions, it was easy to ask one another for help which helped us progress quicker. Some members of our project were more experienced in web development so it was important that we were able to communicate so that the other members could get caught up.

The atmosphere while working together was overall very positive. There was room to joke around, which in some cases could go on for just a second too long. However, we always made sure to bring the topic back to the issue at hand and this allowed for a fun working environment while providing solid results.

6.3 Client feedback

Throughout the duration of the entire project, the client seemed to be excited to find out how modern technology specifically designed for their use case could improve their workflow and the satisfaction of the citizens. During the final retrospective meeting, we demonstrated the final platform. After the demo, the researchers liked the product enough to invite us to submit an abstract for a citizen science-focused conference in October. They also seemed interested in continuing our collaboration and continuing working on the platform until it's ready to be deployed in production. This process involves a future meeting with other citizen science researchers at the University of Twente early in the month of May 2023.

Overall, the clients were particularly excited about the way that we implemented dynamic informed consent. As we have already discussed at length, the platform enables citizens to remain fully anonymous while being able to control their consent in a dynamic and granular fashion.

6.4 Future Work

Firstly, since the scope of the project was rather large, we were not able to fully polish every feature. Some pages contain minor bugs, some wording could be improved and a few optional features are still left to be implemented.

In terms for features, some extra options for the superuser, a more detailed overview page of a survey for the researcher side, general account deletion and export requests would need to be looked at further as well as some smaller features.

Additionally, while many of our citizen pages support a mobile phone view, some pages or items do not fully support this yet. To make the platform fully usable on a mobile device some more work would need to be done on this.

Our platform currently does not support any sort of password recovery and reset through email. While this feature is not necessary for the correct functioning for the platform as a whole, it would be a good thing to have to avoid people having to recreate an account for the platform and possibly losing data that was valuable to them.

Another piece of future work would be do to user testing, specifically on the citizen side on a grander scale. Preferably people that better fit the demographic than those we interviewed during our own user testing, specifically people that suffer from the ailments which the Citizenlab researches. In our own user testing, the user base consisted of our parents. While they fit decently into the age demographic, they might not be representative of the user base of the Citizenlab. Additionally, our user testing was very small scaled which makes it somewhat unreliable by default.

7 Conclusion

This report concludes by answering a question close to the foundation of the Design Project. Did we succeed in completing the design cycle of a software system from start to finish? Starting with an initial preliminary proposal from the Citizenlab we elicited requirements, organized interviews and carefully planned the realization of a novel tool for researchers to work with. This resulted in a project proposal which was, after approval, carried out.

While significant parts of the design were already touched upon while preparing the proposal, codifying technical concepts into a concrete design was our next objective. Any decisions, designs and architectural choices were documented and requirements were adjusted together with the client as needed. Building the platform consisted of multiple phases, each of which produced a deliverable - often a prototype. As the project progressed, these prototypes became more granular and eventually a working product emerged. This product, built using the concepts and principles that arose out of the design phase, was subsequently tested by users. Testing is an important part of the realization of any software system, where ours was no exception.

Further refinement led to the delivery of the end product, which allowed us to reflect on the process as well as the design. The tight timeframe of just ten weeks, only half of which could be spent developing the product, forced us to make choices and balance requirements against delivering a well-functioning system. In hindsight, we believe that the balance was struck correctly as the product is functional and complete to a degree for it to be usable and meet all requirements the clients indicated are a must.

In regards to teamwork, communication with the client and collaboration with our supervisor, no significant roadblocks were encountered. We are satisfied with the dedication of the team and believe that overall good work was delivered. In the end, the clients indicated that their needs were met which would lead to the conclusion that the objective of keeping the client well-involved during the process was achieved.

Considering the former, we believe that we succeeded in completing the design process for a software system and demonstrated our capabilities when it comes to applying theoretical knowledge about software design as well as research skills to a practical assignment. Furthermore, we hope the system we developed may serve as a prototype for future developments, and our design decisions as documented in this report may be taken into account when designing innovative digital tools for research institutions.

References

[1] *Celery*. https://docs.celeryq.de/en/stable/.

[2] Sushant Chamoli and Sarishma. “Docker Security: Architecture, Threat Model, and Best Practices”. In: *Soft Computing: Theories and Applications*. Ed. by Tarun K. Sharma et al. Singapore: Springer Singapore, 2021, pp. 253–263. ISBN: 978-981-16-1696-9.

[3] The Git Community. *Git*. `https://git-scm.com/`.

[4] *Docker*. `https://www.docker.com/`.

[5] *FastAPI*. `https://fastapi.tiangolo.com/`.

[6] *Figma*. `https://figma.com/`.

[7] Python Software Foundation. *Black*. `https://github.com/psf/black`.

[8] Google. *Material Design Guidelines*. `https://m3.material.io/`. 2021.

[9] Maria LaVictoire and Nick Everhart. “A touch screen button size and spacing study with older adults”. In: *SpringerLink* (Jan. 1970). URL: `https://link.springer.com/chapter/10.1007/978-3-642-02707-9_29`.

[10] E.F. Loos et al. “User-friendly websites in the eyes of young and old people”. In: *Proceedings of A transdisciplinary conference organised by COST Action 298 ”Participation in the Broadband Society”, Copenhagen, Denmark, 13th-15th May 2009* (May 2009). URL: `https://dspace.library.uu.nl/handle/1874/40464`.

[11] *MUI*. `https://mui.com/`.

[12] *OAuth 2.0*. `https://oauth.net/2/`.

[13] *OpenAPI Typescript Codegen*. `https://github.com/ferdikoomen/openapi-typescript-codegen`.

[14] *PostgreSQL*. `https://www.postgresql.org/`.

[15] *Python*. `https://python.org/`.

[16] *React*. `https://react.dev/`.

[17] *Redis*. `https://redis.io/`.

[18] *Rome*. `https://github.com/rome/tools`.

[19] Sergiyo Sayago and J. Blat. “About the relevance of accessibility barriers in the everyday interactions of older people with the web: Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)”. In: *ACM Other conferences* (Apr. 2009). URL: `https://dl.acm.org/doi/abs/10.1145/1535654.1535682`.

[20] *SQLAlchemy*. `https://www.sqlalchemy.org/`.

[21] *TypeScript*. `https://www.typescriptlang.org/`.

Appendices

A Diagrams

This appendix contains all used diagrams, on separate pages.

A.1 Data model

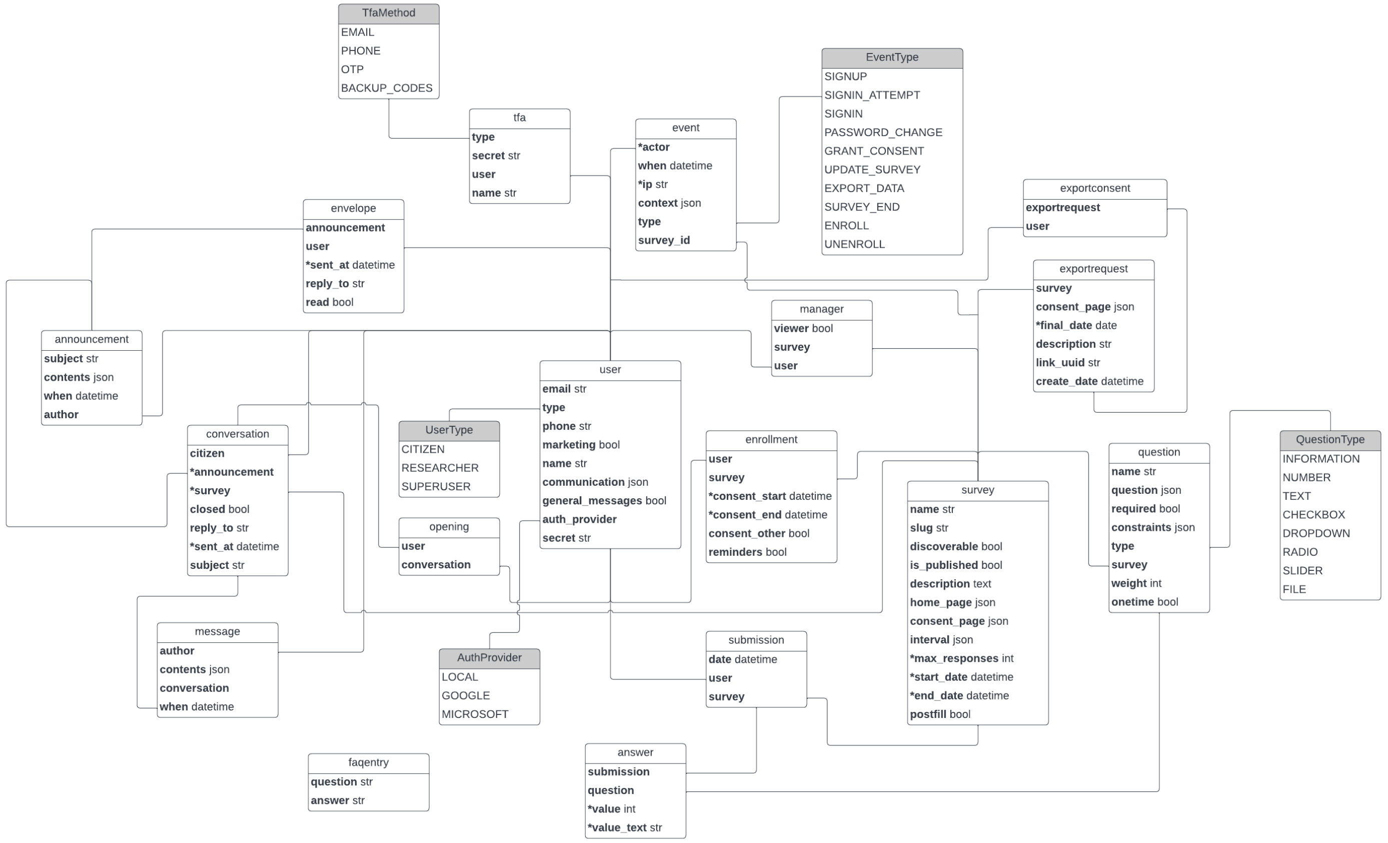


Figure 8: Finalized data model

A.2 Architecture

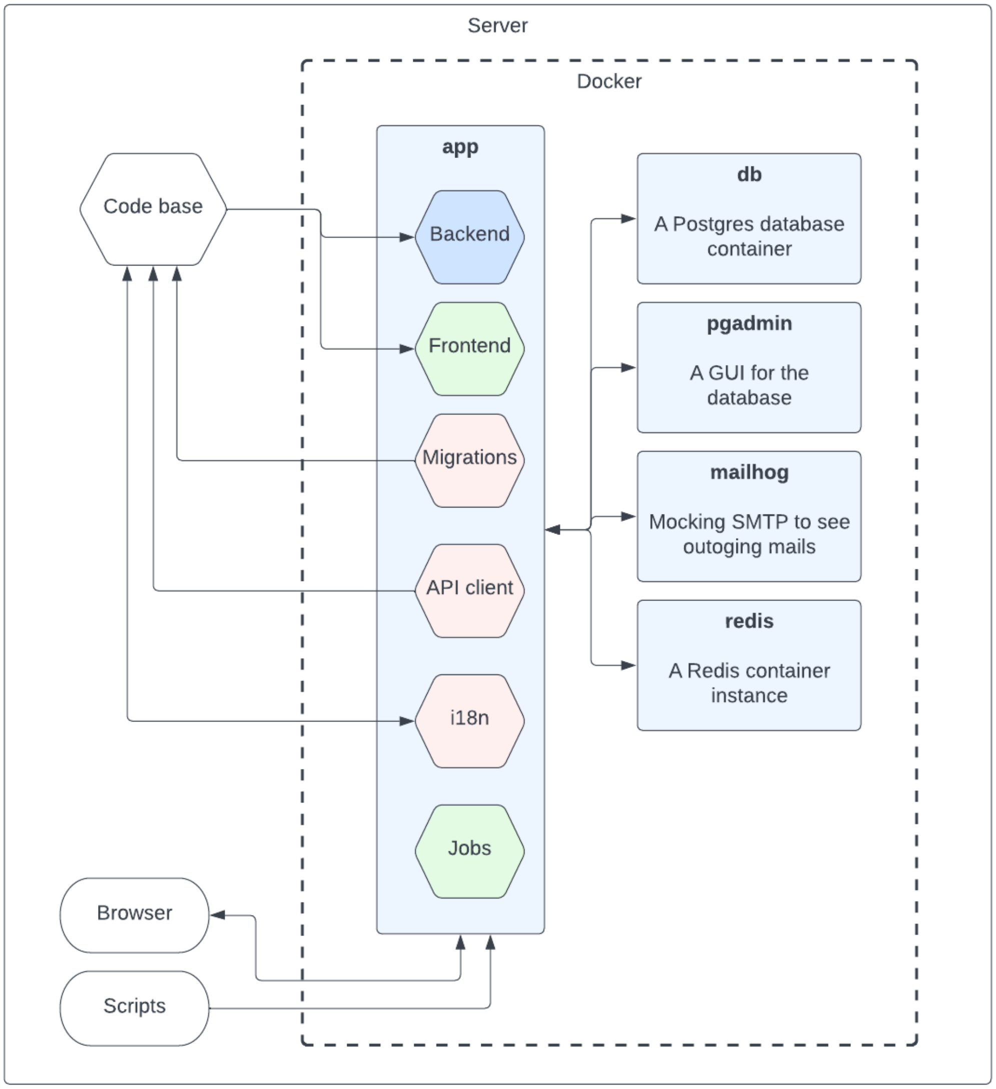


Figure 9: Architecture of the software system, as used during development

A.3 Dynamic Informed Consent procedure

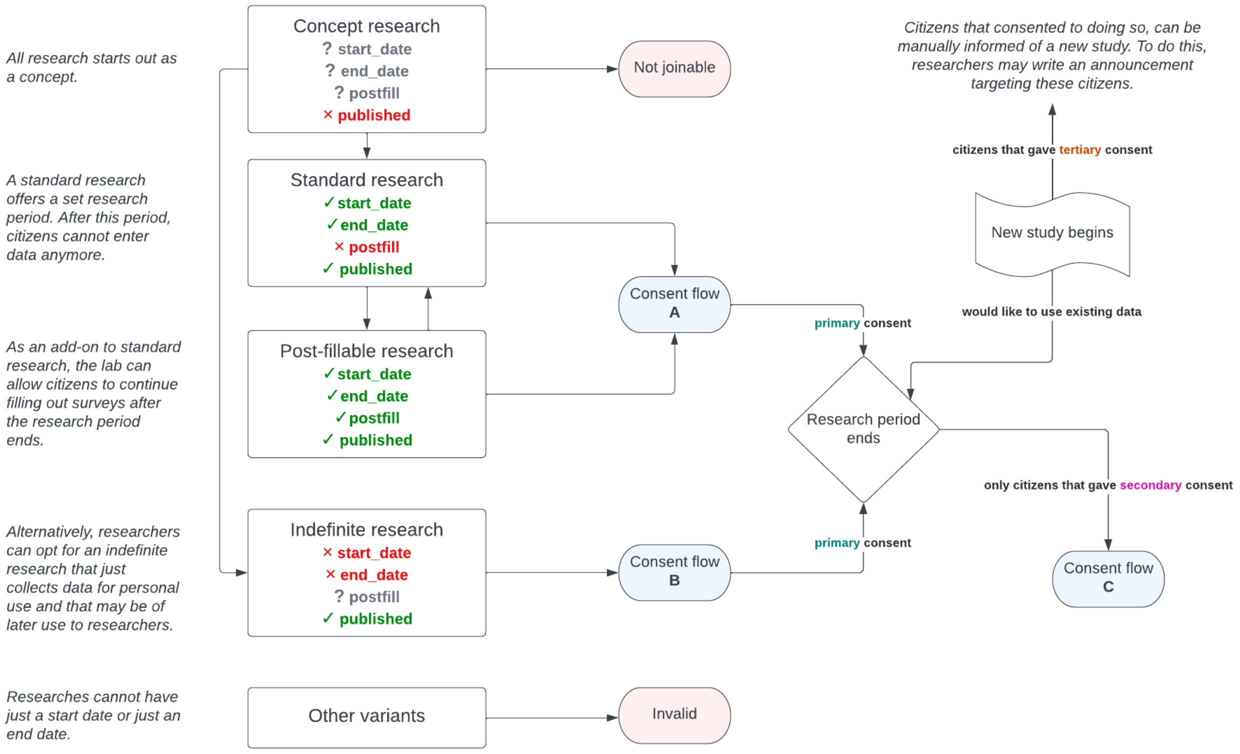
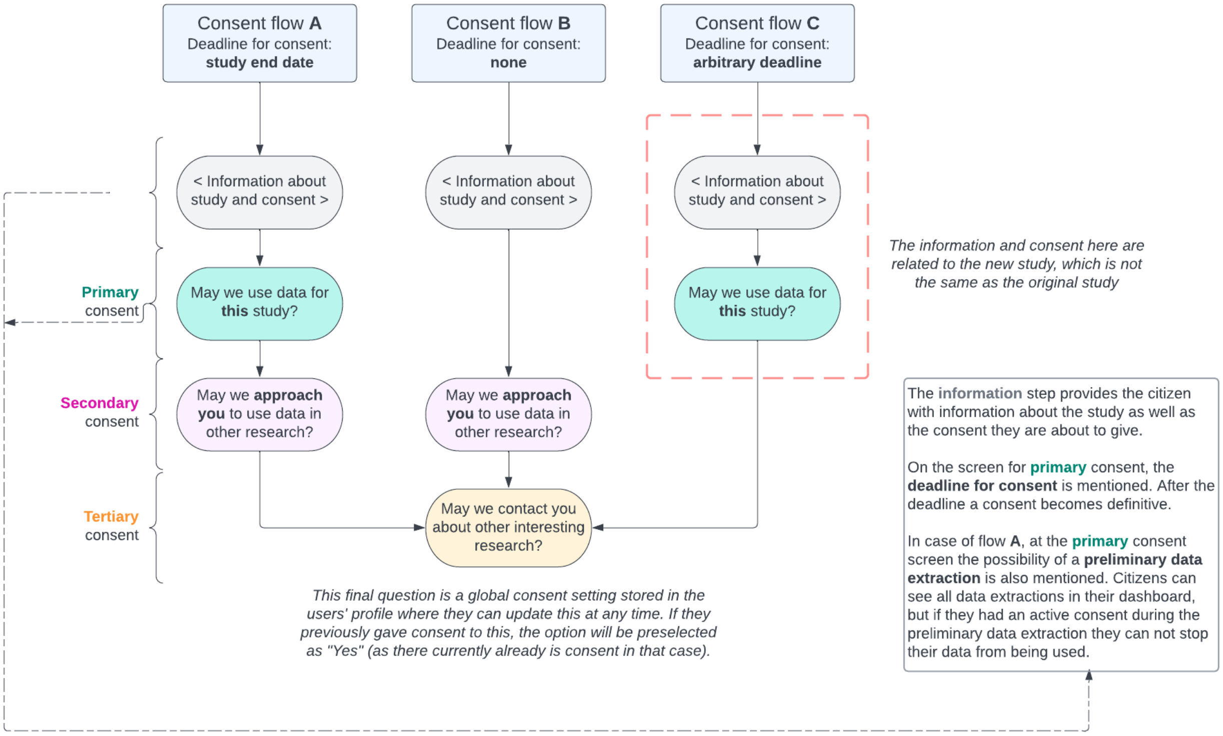

All research starts out as a concept.
Concept research
? start_date
? end_date
? postfill
× published
Not joinable
A standard research offers a set research period. After this period, citizens cannot enter data anymore.
Standard research
✓start_date
✓end_date
× postfill
✓ published
As an add-on to standard research, the lab can allow citizens to continue filling out surveys after the research period ends.
Post-fillable research
✓start_date
✓end_date
✓postfill
✓ published
Consent flow A
primary consent
Alternatively, researchers can opt for an indefinite research that just collects data for personal use and that may be of later use to researchers.
Indefinite research
× start_date
× end_date
? postfill
✓ published
Consent flow B
primary consent
Research period ends
Researches cannot have just a start date or just an end date.
Other variants
Invalid
Citizens that consented to doing so, can be manually informed of a new study. To do this, researchers may write an announcement targeting these citizens.
citizens that gave tertiary consent
New study begins
would like to use existing data
only citizens that gave secondary consent
Consent flow C

Figure 10: Relation between data model and DIC flows


Consent flow **A**
Deadline for consent: **study end date**
Consent flow **B**
Deadline for consent: **none**
Consent flow **C**
Deadline for consent: **arbitrary deadline**
< Information about study and consent >
< Information about study and consent >
< Information about study and consent >
Primary consent
May we use data for **this** study?
May we use data for **this** study?
Secondary consent
May we **approach you** to use data in other research?
May we **approach you** to use data in other research?
Tertiary consent
May we contact you about other interesting research?

The information and consent here are related to the new study, which is not the same as the original study

The **information** step provides the citizen with information about the study as well as the consent they are about to give.

On the screen for **primary** consent, the **deadline for consent** is mentioned. After the deadline a consent becomes definitive.

In case of flow **A**, at the **primary** consent screen the possibility of a **preliminary data extraction** is also mentioned. Citizens can see all data extractions in their dashboard, but if they had an active consent during the preliminary data extraction they can not stop their data from being used.

This final question is a global consent setting stored in the users' profile where they can update this at any time. If they previously gave consent to this, the option will be preselected as "Yes" (as there currently already is consent in that case).

Figure 11: Schematic implementation of DIC flows

A.4 Communication

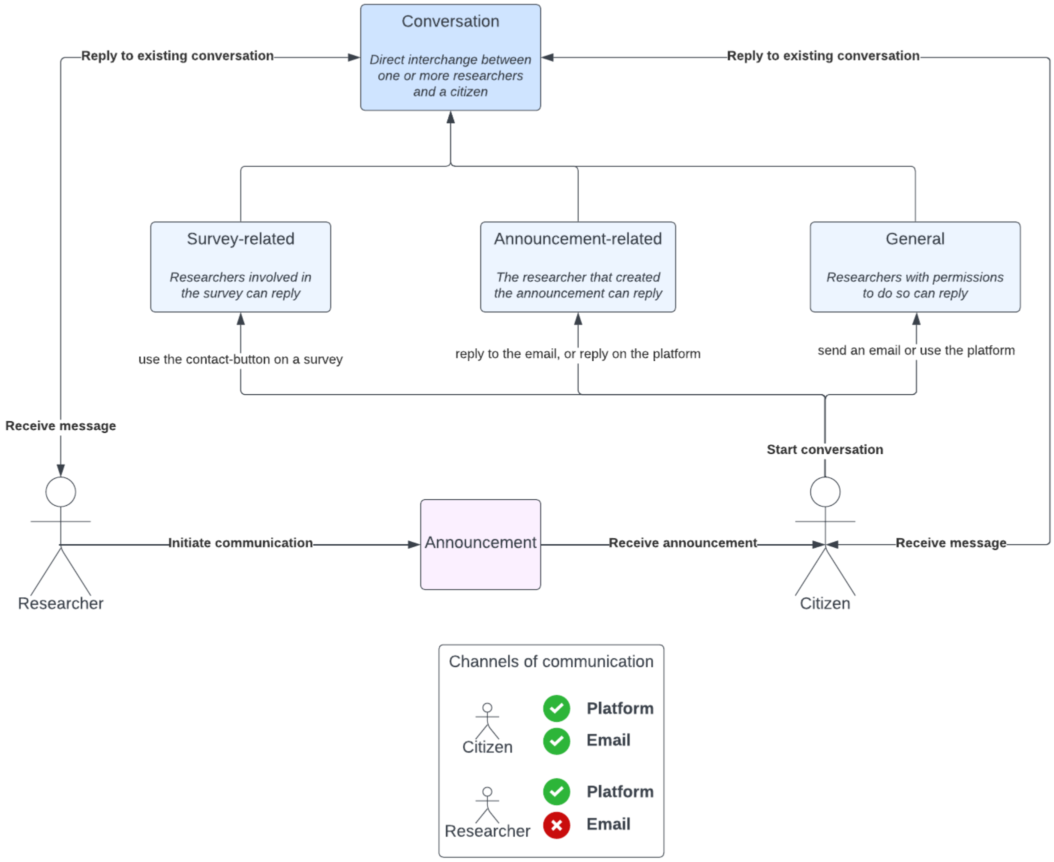

Conversation
Direct interchange between one or more researchers and a citizen
Reply to existing conversation
Reply to existing conversation
Survey-related
Researchers involved in the survey can reply
Announcement-related
The researcher that created the announcement can reply
General
Researchers with permissions to do so can reply
use the contact-button on a survey
reply to the email, or reply on the platform
send an email or use the platform
Receive message
Start conversation
Researcher
Initiate communication
Announcement
Receive announcement
Receive message
Citizen
Channels of communication
Citizen
✓ Platform
✓ Email
Researcher
✓ Platform
✗ Email

Figure 12: Schematic of communication flows

A.5 Interview plan

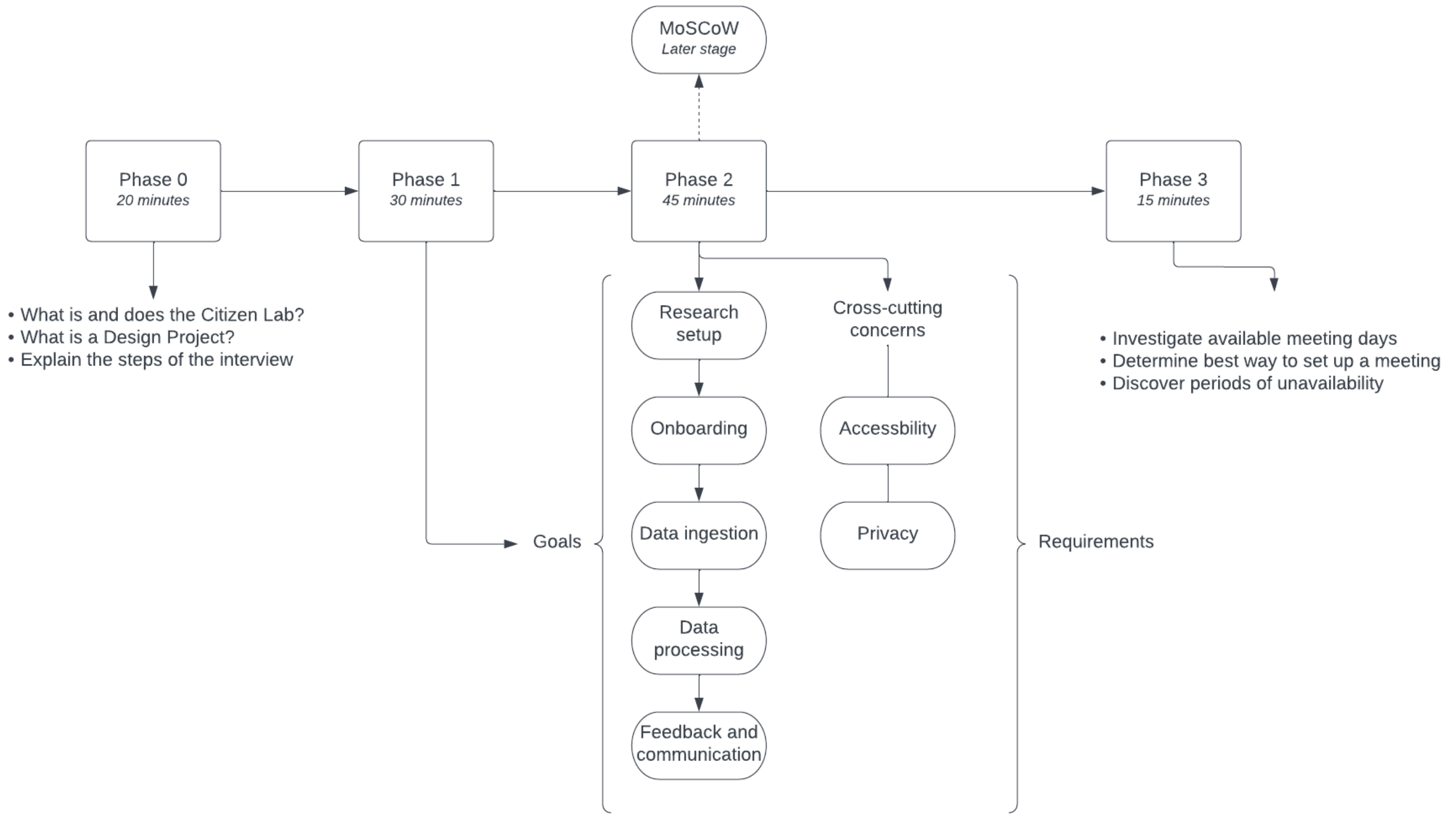

MoSCoW
Later stage
Phase 0
20 minutes
Phase 1
30 minutes
Phase 2
45 minutes
Phase 3
15 minutes
• What is and does the Citizen Lab?
• What is a Design Project?
• Explain the steps of the interview
Goals
Research setup
Onboarding
Data ingestion
Data processing
Feedback and communication
Cross-cutting concerns
Accessbility
Privacy
Requirements
• Investigate available meeting days
• Determine best way to set up a meeting
• Discover periods of unavailability

Figure 13: Schematic overview of the process for the first interview

B Work division

The assignment calls for a work division, which is provided in this appendix.

While it is difficult to pinpoint what every group member exactly contributed to the end product, over the course of the project a rough division of responsibilities established itself.[11] These responsibilities will be outlined below. Responsibilities not mentioned below were a shared concern.

Dawid

- Overall architecture and tooling setup (together with *Mark*)
- Preparing and giving the final presentation (together with *Silas*)
- Work on final report

Jesse

- Poster design (together with *Joep*)
- General interface components
- Question fill flow
- Implementing all question types
- Work on final report

Joep

- Poster design (together with *Jesse*)
- Data visualization in the interface
- Work on final report

Mark

- Leading the client interviews/retrospective meetings (together with *Silas*
- Overall architecture and tooling setup (together with *Dawid*)
- Handling the reflection part of the module (together with *Silas*)
- Implementing the e-mail protocol
- Codifying the DIC procedure
- Work on final report

Silas

- Leading the client interviews/retrospective meetings (together with *Mark*
- Communication module of the platform

[11] Please note that by no means these responsibilities were exclusive to one person, rather they were just the focus of particular group member. In general, the entire project was a collective effort.

- Go to question component
- Handling the reflection part of the module (together with *Mark*)
- Preparing and giving the final presentation (together with *Dawid*)
- Work on final report

C User Testing

C.1 Introduction

Before conducting the test, explain to the user the purpose of the platform, which is research being conducted on the health of citizens. Explain to them that they will try out the website from the perspective of a citizen who will join a survey, fill out a survey, view their own data, and generally navigate around the website. Also explain that user friendliness and giving the citizens a clear idea of why and how their data is going to be used is a core aspect of the platform and they should therefore pay attention whether this is reflected in our design. For any of the steps of the flow, do not tell the user the exact buttons they should click to navigate to the correct page or correctly take a certain action. Let them try out and only help them when they are stuck or very confused about a certain step.

C.2 Flows

C.2.1 Joining a new survey

- Log the user in as a citizen and hand over the controls to them
- Ask the user to change their language to their preferred setting
- Ask the user to try to join the Reuma survey
- Let the user take their time on going through the consent wizard

Questions

- Did the navigation to the different pages feel intuitive?
- Do you feel information was given and explained clearly and did yo feel any information was missing?
- Could you explain in words what you gave consent to concerning your data when joining the Reuma study?

C.2.2 Filling out a survey

- Explain to the user that in this step they will be attempting to fill out the Reuma survey
- Ask the user to navigate to their surveys
- Ask the user to fill in the Reuma survey
- When on question 2, ask them to try to continue to question 3 first. [This will trigger required question]
- When on the submission page, ask the user in what ways they think they would be able to navigate back to question 1 of the survey
- Ask the user to change their answer to question 1 and then hand in the survey

Questions

- Did the navigation for filling out a survey feel natural? Did moving between questions make sense?

- Do you feel the fact that question two is required was clearly displayed? Both in terms of the asterisk showing and the error message which appears when you attempt to skip past.

C.2.3 Viewing a survey

- Explain to the user that in this step they will be attempting to view their data and information on the Reuma survey
- Ask the user to view the Reuma survey
- Ask the user to select two variables and a time frame for their data, but explain to them that the graph won't be very useful when there is only a single data entry
- Ask the user to view their submission history
- Ask the user to view their previous submission
- Ask the user to navigate back to the survey viewing page
- Ask the user to view information about the study
- Ask the user to view their consent settings for the study

Questions

- Did the navigation through the different viewing pages feel natural?
- Did you get a clear overview of your submitted data from these pages?
- Do you feel the information given on the pages, both in terms of your data and your consent is sufficient?

C.2.4 Communication

- Ask the user to navigate to the communication tab
- Ask the user to read the third announcement
- Ask the user to navigate to the messages tab
- Ask the user to select the third conversation
- Ask the user to send any response they want

Questions

- Did the navigation on the communication page feel intuitive?
- Can you explain the difference between announcements and messages?
- How did you experience sending a response?
- Did you notice a difference between messages you read and messages you had not read yet?

C.2.5 Personal settings

- Ask the user to try to navigate to their personal settings
- Ask the user to change their password
- Ask the user to change their name
- Ask the user to revoke all consent
- Ask the user to navigate to their notification settings

- Ask the user to select to only be notified through email

Questions

- Does the navigation to the personal settings page feel intuitive?
- How did you experience changing your password and name?
- Do you feel you understand what revoking all of your consent will do? Could you explain it in words?
- Was the information on notifications clear?

D Ethical considerations for a scientific online health platform

This appendix constitutes an essay that was written as part of a separate course taught alongside the Design Project.

D.1 Introduction

When building a software system that will potentially be used by many users, constraints often come in more variations than purely technical requirements. This is especially the case when that system will be collecting and processing data of a sensitive nature. Over the course of our Design Project, we will design and implement a platform that does exactly that. Therefore, in this essay we aim to reflect and argue about the ethical background of our product.

The Citizenlab is a research group at the University of Twente primarily focused on conducting research with the help of the general public, which consists of many *citizens*.[12] The topics they delve into vary, but for most studies they aim to build a large dataset of data collected by citizens in order to come to sound conclusions. Citizens are asked to self-track a certain metric - some arbitrary examples include weight, pain score, happiness score or how much they drank that day - and report this to the https://www.overleaf.com/project/63e4bb50f60fd5a47d44e910researchers. The researchers will then attempt to put this data to good use, for example by identifying patterns. Participants are sometimes selected based on a pre-existing characteristic or (medical) condition; our clients put forward an example of a study that specifically targeted people with rheumatoid arthritis. Citizens benefit directly and indirectly from participating in the research. First of all, scientific research into themes that play an important role in their life will indirectly benefit them. Secondly, by gaining insight into observations of their own life, citizens may experience a direct benefit because it enables them to recognize patterns themselves and adjust their lifestyle accordingly.

The platform we designed and will be analyzing in this essay consists of a web application that offers functionality to both researchers as well as citizens. Researchers are able to set up surveys: a survey consists of a collection of questions and some settings, such as the desired interval between answers. Citizens can sign up for one or more surveys and will then be able to fill surveys out according to the configured intervals. In close coordination with researchers from the Citizenlab, a *dynamic informed consent procedure* is encoded into the platform, which enables citizens to remain fully in control over their own data.[13] Citizens can view and export their own data to gain insights, and the platform also offers an option for secure and anonymous communication between citizens and researchers.

Our ethical analysis of the platform will start with considerations about the privacy of citizens that use the platform for data collection. Then, we will further explore the ethics of using crowdsourced datasets in scientific research. A few methods to combat ethical issues that may arise out of the usage of the platform will be discussed. Finally, we will draw a conclusion about the overall ethical soundness of the platform.

[12] https://www.topfitcitizenlab.nl/

[13] L. Tauginienė (2021). The Science of Citizen Science: (20) Ethical Challenges and Dynamic Informed Consent.

D.2 Ethical considerations about privacy

Possibly the most prominent ethical consideration relevant to our product is the question of privacy, so we consider this first and foremost. The answer to the question if collecting possibly sensitive data from many people is ethical is very complex. Therefore, we will first specify some prerequisites that need to be met in order to be able to consider the data collection ethical. Then, we will try to narrow down the actual consideration to be made in deciding if data collection is acceptable. Finally, we will impose an ethical question for the platform which arises from this consideration.

It is easy to define scenarios in which data collection would be considered unethical by most. Secretly collecting the data, for example, would most likely be unethical. Voluntary and informed consent from participants is therefore a must. Treating the data uncarefully, such as storing it in an unsafe place or publishing personal information of participants is unethical too. In general, we would argue that basic safeguards must be in place when using data for scientific research. Some of those principles are already present as legal requirements, especially in more recent legislation.[14] Other principles are often codified in ethical guidelines from institutions performing research.[15] Since violations of those well-established ethical principles are clear ethical missteps, they deserve no further attention in this paper.

In what other circumstances would we consider the collection of data to be unethical? A useful viewpoint to analyze the ethics of an activity is to investigate whether any harm could potentially be brought upon those involved and compare this to which good the activity can bring. The advantages of data collection are quite obvious: they contribute to scientific research, which in turn helps out everyone. The downsides of the data collection itself, considering the data is only used for its intended purpose, are pretty innocent too, though we will discuss some objections to this in the next section. There do, however, exist ways in which the activity of data collection can inflict harm on our citizens. Should their private data leak and become public, for example, this would lead to damage that is hard to oversee or repair. Fortunately, due to aforementioned safeguards, every individual involved in the study may relatively safely assume such things do not happen. However, we still need to take into account that these events, however unlikely, *can* occur and ensure the positive effects of producing the research outweigh the negative effects of harm the data collection causes. Obviously, the gravity of the negative effects need to be put into the right perspective: the vast majority of studies will not suffer from data leaks and the subsequent negative effects. Be that as it may, collecting and then keeping large amounts of personal data without a reason would be considered unethical. The benefit in that case is nothing, while the risk is nonzero. We argue that research carried out by the Citizenlab is far from useless, yet one must wonder when a study is useful enough to warrant the associated risks.

D.3 Crowdsourced datasets for research

A fairly hidden, yet truly important ethical consideration that also plays a role is about using a large crowdsourced dataset for the purposes of scientific research. After all, while the research in large part exists purely to serve society, the interests of institutions pursuing this research are nearly always still commercial at their core. To successfully carry out their studies, the Citizenlab asks citizens to provide a bit of effort. While the amount of effort required by a single citizen may be considered negligible, the total amount of human effort that goes into a single study will be of significance, due to the large number of participants. Is it unethical to pursue the goal of collecting small bits of data from many people and then benefit from this, even if a large part of this benefit also constitutes “doing good for society”?

A good starting point for exploring this question may be to try and answer the question why it would *not* be ethical to build and then subsequently use large datasets of crowdsourced data in research. At first glance, a natural response to that question may be to feel like it could never be wrong to use data that was collected with permission to do good for society. The dilemma, however, may very well not be as simple as that. From a utilitarian perspective it is thought to be a good thing when many tiny contributions can be converted into

[14] G. Chassang (2017). The impact of the EU general data protection regulation on scientific research.

[15] University of Twente (2019). Code of Ethics University of Twente. Available at www.utwente.nl.

a larger “unit” that is worth more than the individual contributions we started with. Extending that line of thought, as long as the value of a study outweighs the aggregate effort put in by our citizens, the collection of data is ethically justified. On the other hand, from a deontological viewpoint one might consider making someone work primarily in your interest without offering any form of compensation is wrong no matter the circumstances.[16] The fact that the amount of effort is relatively small does not play a role if you follow this line of reasoning.

Parallels can also be drawn to the buzzword of *big data* - the phenomenon of large companies collecting (small bits of) data from millions or billions of users. While the goals of these companies are usually purely commercial, the practice of data collection is strikingly similar. Many have argued before that many practices of the big data industry should be dismissed on ethical grounds.[17] Another hotly debated topic is the use of open source code for the training of AI models.[18] The connection to our platform in that regard is seemingly a bit weaker, but proves significant after all: this also concerns an *end product* (in our case scientific research by the Citizenlab) which is derived from resources for which no *compensation* was offered (in our case effort by citizens) on a very *large scale*, resulting in a very low intrinsic value for each unit of effort offered but a substantial intrinsic value for all effort when composed. So, if these concerns are so prevalent when it comes to commercial data mining, should they not also apply to the collection of data for scientific purposes?

D.4 Responding to ethical concerns

After establishing ethical concerns arising out of using the designed platform, it makes sense to incorporate remedies for those concerns into the design. In this section, we aim to outline our responses to raised ethical concerns as well as their costs and benefits. First, we will argue that technical security measures can guarantee privacy to a sufficient degree in order to ensure collecting personal data can be considered an ethical activity. Then, we will highlight how an informed consent procedure can help ensure researchers make ethical use of human resources.

It is impossible to guarantee users of our platform a data breach will never occur. Due to the high impact of such an event, however, it is of high importance that the risk of a data breach is small enough to warrant collecting user data for research purposes. It is important to also realize data breaches come in many different forms and sizes; they do not solely arise from cybercriminals illegally entering digital systems. Research into (digital) theft and disclosure of medical data records shows that a significant amount of breaches take place by physical means, such as the theft or disappearance of hardware, inside attacks or personnel losing data carriers such as thumb drives.[19] Smaller and simpler internal breaches, such as wrongly assigned permissions also belong to the possibilities when it comes to a platform that primarily processes data. Upon implementation, technical measures to prevent data breaches were granted special attention. However, data protection knows many dimensions of which prevention is only one. Detection of a (possible) breach, and storing only a necessary amount of data are additional measures that should be taken. By taking these measures into account throughout the entire design cycle and realizing an implementation with a secure foundation, we believe enough care is taken to ethically source data of the general public using the platform.

Another ethical concern raised regards the use of human resources for the production of scientific research. Is it ethical to use a (in total) large amount of human effort to pursue a goal of own interest? Analogies suggest that the only difference between a social media platform collecting user data and the Citizenlab performing research is that research generally does not primarily satisfy commercial interests. However, we do believe there exists an important mechanism to justify data collection for research purposes. This mechanism is the (dynamic) informed consent procedure embedded into the platform. Exactly this procedure sets responsible data collection for research apart from the often implicit data collection practiced by the big data industry. By being transparent and upfront to participants about *why* and *how* we collect and use their data, we

[16] Examples of different ethical viewpoints taken from the lectures; topic II

[17] K. Martin (2015). Ethical Issues in the Big Data Industry.

[18] A. Al-Kaswan (2023). The (Ab) use of Open Source Code to Train Large Language Models.

[19] A. Hussain Seh (2020). Healthcare Data Breaches: Insights and Implications.

can ensure that the decision to participate is based on the right grounds. Citizens actively and consciously consent to invest a little of their own effort and give out a little of their private information only because they support the goal: aggregating information to make it significantly more useful and serve a good purpose. The dynamic informed consent flow is a primary and proud feature of the platform and the embedding of ethically sound principles will hopefully contribute to using crowdsourced user data in an ethically acceptable manner.

D.5 Conclusion

Even though building a functioning online platform to collect research data could be considered an accomplishment based on purely technical skills, we believe that taking a step back and reminding ourselves that ethical considerations are at play as well is an important part of the design cycle.

As we have shown, ethical questions can be imposed from numerous different angles and viewpoints. Some questions are more obvious than others: privacy considerations are natural to consider when building such a platform, while the unintentional exploitation of participants is a more hidden issue. All questions have in common that answers are complex and moreover different based on individual viewpoints. As the designers and developers of the platform, we are not the only ones bearing this responsibility. We do believe, however, that by having a deeper understanding of the matter and incorporating ethical considerations and safeguards into the product, the end product will become better.

Examples of enhancing the product with an ethical understanding can be found in our responses to the ethical concerns. By investing a proper amount of resources into security and aiming for a secure platform by design, we stimulate a more ethical manner of processing user data. With our gained knowledge of the importance of informed consent, we were able to build a better consent flow which in turn helps researchers conduct their experiments in an ethically sound manner. Taking all of the above into account, we believe to have developed a platform that fosters ethical behavior and hence can contribute to ethical scientific research.

E Interview questions

E.1 Goal questions

These general goal questions were asked in part I of the interview.[20]

- What problem are you trying to solve?
- Can you briefly describe your current solution?
- Is your current solution lacking in some areas?
- Can you briefly describe the flow of a normal research study?
- How do you find participants?
- What about privacy?
- What is the user base like (age / demographic)?
- Why do users participate in a study?
- What are your goals regarding the administrative part of the platform?
- What sort of things do you need to be able to do?
- What sort of information do you (as a researcher) need to be able to view?

[20] For the interview plan, see appendix A.5

- What information should users be able to view themselves?
- What information needs to be communicated to the user?
- What are your goals regarding accessibility of the platform?
- What should we keep in mind for how a client interacts with the platform?

E.2 Requirement questions

The questions below were primarily used to elicit requirements.

Research setup

- What kind of input methods would be required (textbox, slider, text, checkboxes)?
- Should the questionnaire be able to support images in the questions?
- Should the questionnaires be able to be edited after publication?
  Basic editing: changing question message
  Advanced editing: adding/removing entire questions/question types
- What can be the length of a research study?
  Must this be variable?
- At what frequency will users be able to answer?
- Must this be variable?
- What happens if a user doesn't respond for one or multiple data points?
- Is there a need to bulk import existing study participants from your existing system?
- Should the result from a questionnaire be visualized in some sort of graph?
- Are there multiple research groups making use of the platform?
  Should they have different permissions?

Onboarding process

- How does a user get recruited/registered into the system?
- Do you perform identity verification?
  How does that process currently look like?

Data ingestion

- Will the user be reminded it's time to fill out the survey?
  How? E-mail, text? Something else?
- Do you have any vision regarding the login process of the user?
- What sort of login method would be desirable for the application?
  Federated sign-in using an official platform like DigiD?
  Social media sign-in?
- Does a user need to authenticate for every questionnaire?

Data processing

- Do you want to be able to remove certain individuals from research?
- Do you need any built-in data analysis features?
- Do you need to be able to export the data in some sort of format?

Feedback to the user

- Do you want to provide feedback to the users, and if yes what sort?
  What kind of feedback? Through the site / emails / advice?
- Do you need to provide feedback to an individual user (anonymously)?
- Do you need to provide feedback to a group based on results obtained?
- Do you want the feedback to be two-way (users can reply to the feedback)?

Accessibility

- What kind of medical conditions do you expect your users to have?
  Try to discover if these conditions might impact the use of a potential application
- Would the platform require some sort of helper/tutorial mode?
- Do you believe it is valuable for the platform to be available on different devices?
  Such as a mobile application?

Privacy

- Who owns the data: users or study?
- Can data be shared amongst studies?
- When you sign up do you opt in by default?
- What does your informed consent process look like?
- What information of the users and their responses will the researchers be able to view?
- Is there a requirement to purge data of inactive users after a certain period of time?
- How long does data need to be kept after a study?
- When data needs to be deleted, is it enough to cut ties to any personal information?

F Proposal

The approved proposal is attached verbatim as an appendix to this report.

Design Project proposal - Citizenlab

Building a health platform for citizen science

Dawid Kulikowski (s2472910)
Jesse Snoijer (s2572362)
Mark Boom (s2552469)
Silas de Graaf (s2220032)
Joep Vorage (s2172968)

February 2023

Table of Contents

1 Introduction

The purpose of this proposal is to outline a software system meant to assist researchers of the Citizenlab of the University of Twente in their research. This will be done by investigating the current problem as well as presenting requirements for a solution. Accompanied by a plan for further research, a risk assessment and potential timeline, this proposal serves as a good basis for building the envisioned solution.

1.1 The Citizenlab

The Citizenlab is an organization which performs research together with people from society ("citizens"). In many studies, gathering interesting data with the help of citizens is a primary goal. Research studies are developed hand-in-hand with the citizens: they are encouraged to actively participate from start to end. The Citizenlab gathers the data needed mostly through the use of surveys which are filled out by people who often suffer from a medical condition. At the moment of writing, the medical condition the lab is focusing on is rheumatoid arthritis. Citizens with rheumatoid arthritis are asked to be part of a research about the condition. They then are asked to fill out a daily survey that gathers data with interest to their condition.

The aim of Citizenlab is to involve all stakeholders in this process. When asked to perform research, the Citizenlab starts with creating a research question and deciding on the data needed. Then they organize a meeting with the citizens and involve them in the process of creating questions. Doing so helps citizens be more involved with the research. Another way Citizenlab tries to involve citizens is by giving them insights into their own data. This way citizens are able to track several aspects about their health over time and draw conclusions specific to themselves.

1.2 Designing a software system

In this document we will propose a web application that will fit the needs of the researchers of the Citizenlab when it comes to data collection. The target audience consists of citizens with a medical condition where we focus on citizens with rheumatoid arthritis. The process was kicked off by identifying the exact problem. Since currently another external platform is used by the lab, any difficulties with the current solution were mapped out too. Then, this proposal lists the different ways research was performed (to draft this proposal) and any further (user) research required to complete the system. Thereafter, a solution is presented to solve the identified problem. This will be in the form of a web application developed by us. We also provide a timeline for the expected duration of this project. In this timeline, a few milestones are set. An assessment of the possible risks during and after developing the proposed solution can be found at the end of the proposal.

A comprehensive list of requirements as well as an activity diagram have been composed and included in the proposal.

1.3 Glossary

citizen (with rheumatoid arthritis), participant	user filling in surveys
client(s)	researchers from the Citizenlab who are involved in the project
health platform	the product / (web) application we create
researcher	user creating and managing surveys
(end) users	citizens & researchers
dynamic informed consent	extended "informed consent" procedure which provides the option to control consent on a very granular level (select exact data points) and allows for consent adjustment as often as desired

2 Problem

2.1 Goal

The goal of the Citizenlab is to involve citizens in the research about their own medical condition. Citizens will collect data about their health, to either gain more insight or to test the effects of certain interventions. This is done in a scientifically sound manner, with groups of varying size. For instance, people with rheumatic conditions can record both fatigue scores and physical activity for a month to identify a pattern, or they can record pain scores a month before and after starting a vegetarian diet to measure an effect. Or an individual with Parkinson's disease who would like to better understand the effects of the time of day in which medication is administered on their stiffness. Such data can be collected for personal insight or analyses and shared with researchers for scientific projects.

2.2 Current solution

The Citizenlab currently collects all of its data through the use of an external tool. Once the research itself has been designed and configured on the platform, they will recruit people through social media as well as other places like hospitals where they are likely to find their target audience. The citizens will be recruited through an URL, where they have to sign up which requires an email verification step. Once they are signed up, they are coupled to a survey which they can then fill out once in every predetermined time period (this is often daily). The current platform supports email and SMS reminders to help the citizens remember to fill out the survey, which the researchers found to be a handy feature. Additionally, the external tool supports dynamic informed consent which is of great importance to researchers. This approach, however, does come with multiple issues when it comes to the ease of use for both researchers and citizens. First and foremost, data collection taking place on another platform is something that the researchers do not prefer in general. They prefer citizens being on a platform they provide. This gives them more control over the flow and collection of data as well as the communication with citizens. Our clients reported that during specific occasions, like for example requesting the emails of their citizens so that they can be offered to be part of future research, they would need to contact the administrative team of this external tool for some specific request which could be a long and tedious process. They report that they would prefer being able to communicate with the citizens anonymously in some way so that the citizens can simply give them access to or provide certain data. This, of course, all while making sure that the citizens consent with each step along the way. Finally, our clients mentioned that the data of this external tool would be exportable only to the JSON format which they find bothersome to process compared to their more preferred format, which is CSV.

Our clients reported multiple difficulties citizens are having as well. They noted that in their experience, the external tool seems like it is meant for a younger target audience. The interface in general can be complicated to use for citizens who are not tech-savvy. Many citizens would drop out already at the process of account creation when they would have to verify their account through email. More specific things were mentioned too, like color contrast and needing to scroll on certain pages. People with impaired vision have difficulty reading items on a web page when it is gray on white, and people who have some sort of impairment when it comes to their fine motorics find it difficult and sometimes painful to produce the motion necessary for scrolling through web pages.

Aside from difficulties in the citizen user experience, citizens also reported on wanting more flexible feedback. The external tool allows citizens to view their own data. The tool supports citizens to view their collected data through simple graphs over time on their website, as well as the option to export one's data. This viewing of data is of importance to the citizens as it helps them get a picture of how their health has been doing over time. However, the tool is not as flexible in this as citizens would like it to be. Citizens report wanting to be able to relate different variables and view them coupled in a graph. A citizen who has been keeping track of both their energy levels as well as the amount of exercise they have been doing would like to be able to view them in a single graph to view how they relate to one another. Additionally, citizens wish to be able to zoom in on specific sections of their data, wishing to view only a specific month for example. Unfortunately, the current tool only supports these simple straightforward graphs of single variables which can make it difficult for citizens to review their own data visually.

To summarize, researchers would prefer having the citizens on their own platform so that they can provide information to the citizens better as well as communicate with them more flexibly. Additionally, the exporting format is currently not to their liking. On the side of the citizens, ease of use is a great concern considering the target audience of the Citizenlab which is currently lacking. Additionally, citizens value viewing their own data and believe this is something that could be expanded on to be more flexible than it is currently.

2.3 Importance

Finding a solution that fits the wishes of the researchers is important. Citizen science is a very exciting upcoming field due to the potential of directly connecting many citizens with science while at the same time having a profound positive impact on the lives of these citizens. Without a solid digital platform to work from, researchers from the Citizenlab spend a lot of time on practical matters instead of research. Therefore, we do believe this is a useful problem to work on.

3 Research

A solid computer system can only be implemented if the goals and requirements of the target users are clearly identified. In this section, our target group will be defined. Furthermore the setup of user research is illustrated. Accessibility issues are investigated using literature research.

3.1 Target group

The platforms' users can easily be subdivided into two main groups. The first group consists of researchers that set up studies, collaborate, invite citizens and eventually use the system to export data to process. The second group consists of citizens: practically any member of society that in one way or the other causes data to enter the system. They have their own journey of getting registered and then adding data. After doing so, they can then decide to share this data on a voluntary basis with research studies.

3.1.1 Researchers

This platform will initially be developed for the Citizenlab of the University of Twente. A small group of researchers is involved with citizen science and will be using the system for their studies. The researchers are academics with varying backgrounds. While having the skills to deal with various digital systems, care must be taken to abstract technical details such as data modeling away from the user. A functional system with an explicit focus on being able to set up research studies as easily as possible is desired. This target group will be explicitly involved in the development of the system.

3.1.2 Citizens

The citizens together form a very large and very broad target group. It's debatable whether this can even be considered a target group due to the fact that in principle any person is inherently part of this group.

However, having concluded that in theory anyone can participate as a member of this group, it is still possible to take the reality into consideration and investigate if there are specific traits among the subgroups of all citizens that are most likely to use the platform (at least initially). A few characteristics can be identified:

- Most participants live in the Netherlands and speak Dutch;
- Many participants are of older age;

- Often, participants have below-average technical skills; [1]

Aside from these general characteristics, much of the research performed focuses around certain diseases or disabilities.[2] It is therefore of value to try to cater to the specific needs of the people subject to these diseases.

3.2 User research

The main source of information regarding requirements and goals of the platform originates from the researchers that will eventually use the product. To properly ensure their involvement in the design process, a process was constructed that involves them at different points of the design cycle.

3.2.1 Requirements interview

To start off, an initial interview was organized with the purpose of requirements and goal discovery. All requirements in this proposal are based on this interview, supplemented by some research.

The two hour interview consisted of four phases. Two of which were for practical purposes, while the other two provided the goals and concrete requirements respectively.

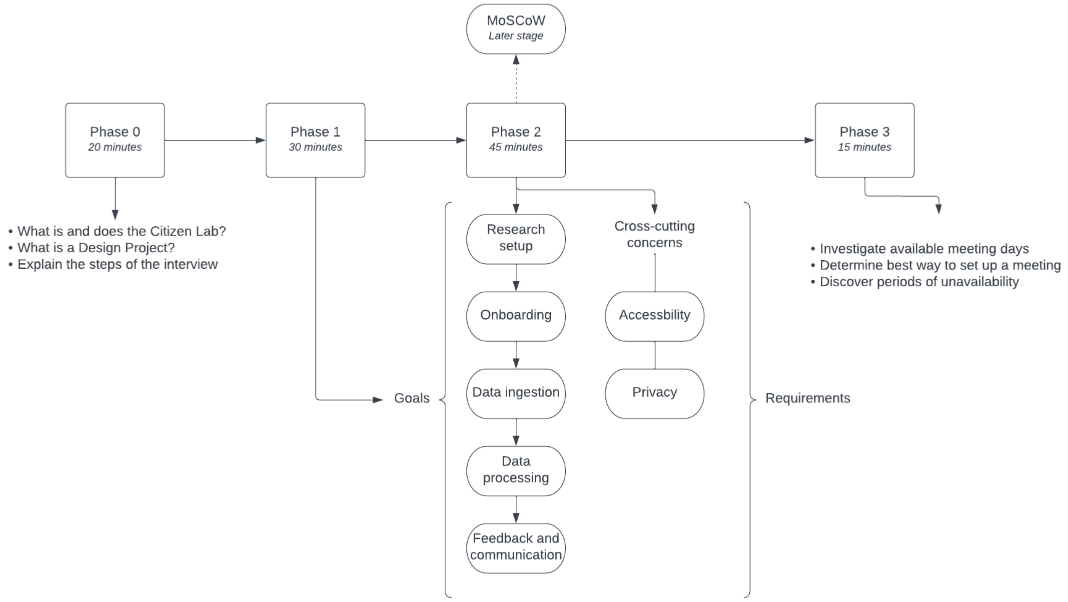

MoSCoW
Later stage
Phase 0
20 minutes
Phase 1
30 minutes
Phase 2
45 minutes
Phase 3
15 minutes
• What is and does the Citizen Lab?
• What is a Design Project?
• Explain the steps of the interview
Research setup
Cross-cutting concerns
Onboarding
Accessbility
Data ingestion
Privacy
Data processing
Feedback and communication
Goals
Requirements
• Investigate available meeting days
• Determine best way to set up a meeting
• Discover periods of unavailability

Figure 1: Schematic interview setup

At first, very broad questions were asked to the interviewees. While not particularly useful for requirements elicitation, these questions were meant to discover the main goals that the platform should achieve. Goals were categorized in real time, for use later in the interview. Categories were created by the interviewers

[1] The main source of this information is an estimation by the researchers. It's hard to express the digital literacy of such a large group in numbers, but the fact that warning signs are present should lead to the conclusion that ease-of-use is of high importance in the design process.

[2] Rheumatoid arthritis, for example

prior to the interview, based on the initial project proposal as well as earlier informal conversations with the interviewees.

In the next phase requirements were elicited based on more direct questions. Questions were asked per category to maintain a proper flow within the interview. Most questions were prepared beforehand, though some questions originated from the interviewee statements at an earlier stage of the interview.

3.2.2 Prototypes

Most user research in this project will be conducted with the help of prototypes. As progress on the final product progresses, prototypes will become more sophisticated. The following prototypes will be created:

- **Low fidelity prototype**
  This initial prototype will consist of a few simple (digital) drawings of an interface. It will allow us to introduce the clients to our view of the problem for the first time and serves as a good basis to discuss if all requirements are included.
- **High fidelity prototype**
  The next prototype will be the basis for the actual user interface that will be built for the product. There is more focus on how interactions work and there is more room for discussion about smaller user interface details.
- **Minimum Viable Product (MVP)**
  The MVP will be the very first iteration of the product. All requirements and principles such as presented in this proposal and the earlier prototypes will be functional with both a working interface and a backend. The clients can use the product for the first time, which allows them to provide effective feedback to shape the final product.

3.2.3 Interviews

Throughout the project, the clients will attend interviews to provide feedback and express any further wishes. Interviews will be planned in close coordination with the prototypes as they are developed. This allows us to work from a viewpoint of very general concepts and little detail to be able to discuss minor details directly with the clients towards the end of the project. See 5 for more details.

3.3 Literature research

To assist us in making initial decisions about, among other things, accessibility issues we performed literature research. We mainly considered research done into the accessibility of websites for elderly people as well as people that are physically impaired in some manner. This research shows that older citizens need more time and follow a different reading pattern [2]. So in order to build a user-friendly web application, we would prefer any pages visited by citizens to be as simple as possible. This can be achieved by sticking to a minimum when it comes to interface elements: for example by only providing users with buttons they are very likely to use and avoiding an overload of options to choose from or distinguish between. Another study presented a table with a manageable and robust set of guidelines for designing and evaluating age-friendly websites. The guidelines were supported by published literature and have been subjected to several stages of expert and user validation, which should provide some assurance of their validity [5]. Abiding by the guidelines specified in this table should increase chances of a user-friendly experience for our target group.

Specific research has also been performed into the user experience of older people. A study performed in 2009 [1] contains a study on the button size preference of older people interacting with a touch screen device. In this study, it was determined that the minimum button size for this age group is 1.25 cm x 1.25 cm, which should be considered when designing a phone or tablet application for our target group.

Older people have also been found to prefer words over icons. Another research performed in 2009 which studied several hundreds of elderly individuals found that many reported on prioritizing textual information. They reported that icons can sometimes be unclear when it comes to their meaning whereas they can feel confident in their interpretation of the meaning of textual parts of a web page [3].

In terms of physical impairment, one article [4] discusses a couple of methods to alleviate some of the difficulties that people with physical impairments might face when browsing the web. This includes avoiding double clicking, avoiding intensive scrolling behavior, and having the website be operable through a single medium where possible. These mediums in the case of our platform are mouse, keyboard and touchscreen. Actions that require a long, continuously held movement (such as clicking and dragging an object with a mouse) should also be avoided where possible.

4 Solution

To achieve the goals described in section 2, a platform should be developed that builds on the strong points from the current solution while also tackling its associated problems. This section describes the proposed solution in detail.

4.1 Requirements and activity diagram

From interviewing the client, a list of requirements has been composed in Appendix A. It is important to mention that while the list has been made as complete as possible it is inherent to the agile development process that requirements are to be added and altered over the duration of the project. Hence, these are not necessarily the final requirements. A weekly internal reflection on what was developed ensures requirements can be responsibly updated, added or removed.

To illustrate the flows a user can follow a bit better in a visual way, an activity diagram has been included in Appendix B.

4.2 Differences with current solution

To understand the necessity of building the platform it is important to highlight the core differences between the gathered requirements and functionalities the current solution provides.

- The platform will facilitate creating and editing surveys as a researcher which is not possible in the current solution.
- The platform will be tailored specifically to the Citizenlab's target group, mostly elderly people, while the current solution is mainly targeted towards young people.
- The platform will allow for anonymous communication between researchers and citizens while the current solution does not have any way for researchers to contact the people participating in their surveys.

4.3 Value of solution

The main value that the platform will provide for researchers is independence from any third-parties. The ability to create, edit and extract data from surveys makes it possible to work more productively. It will also remove the risk of the third party being unavailable or unwilling to publish the proposed survey.

Another very valuable characteristic of the platform will be that it caters to the core business of the researchers. Every aspect of it will be created with the interest of researchers and citizens in mind. As an example, the data to be extracted from surveys can be formatted to fit the preferred data analysis tool used by the Citizenlab. Another helpful feature of the platform could be its ability to enforce certain academic rules, such as prohibiting data extraction before a set date.

There is also a lot of added value for the citizens making use of the platform. Mainly its accessibility, which will also be considered during every aspect of the design process. Not only that, because the larger part of citizens are expected to be less tech-savvy than an average user, special care will be taken when designing the registration and login flow to ensure its simplicity.

4.4 Impact on target group

The platform will have the potential to have a big impact on the target groups. Firstly, the researchers' workflow will be more streamlined as they now have access to a platform designed specifically around their needs. This allows them to fully focus on the study and the questions they want answered, and ensures practicalities like reaching citizens does not distract from their work.

Citizens are now able to take full ownership of their data. More than ever before, they get to decide what data they collect and how it gets used in research. Lots of flexibility in granting or retracting consent allows citizens to feel very comfortable using the platform to track their own health. Involved researchers reported that some citizens feel more valued knowing they are actively contributing to scientific research. By combining that with the fact the platform will allow citizens to also get value from participating (such as new insights into their health), the platform will feature a powerful combination between it's users and science.

5 Timeline

The timeline below was constructed to roughly estimate what work should happen in which weeks. Milestones include progressive updates to the product as well as retrospective meetings. Clients will be invited to the retrospective meetings.

Week	Milestones
Week 1	Intro meeting, Acquire supervisor
Week 2	Requirements interview
Week 3	Design proposal, Lo-Fi prototype, Client meeting
Week 4	
Week 5	System design, Hi-Fi prototype, Retrospective
Week 6	User testing
Week 7	MVP, Retrospective
Week 8	
Week 9	Final product, Report, Exit interview
Week 10	Poster, Final meeting

Figure 2: Milestone timeline

6 Risk analysis

Building any software system comes with certain risks. During the design process, we have to take into consideration any risks for the citizen and the researchers that are going to end up using the system we develop. What if an engineering mistake leads to user data leaking to the public? How much bad luck is required to derail the entire planning of constructing the system? For this project, risks can be divided into a few main categories: risks related to privacy and security, risks related to the process of building the system and risks related to the operation of the system itself. This section explains how risks are discovered and managed. As the system deals with sensitive health-related data, we mostly focus on issues related to storing and handling health data, notably any issues that can cause bodily harm to the users and privacy concerns.

6.1 Privacy and security risks

After the system is built, the main risks are related to user privacy and overall security. The storage of health data presents unique problems. Any data leak can contain strictly confidential information ranging from a personal diary to blood test results. Anonymous users can also have their contact information publicized which is clearly undesirable. Such information should not be involuntarily publicized and can cause a variety of issues for both the citizens and the researchers. The citizens may become victims of identity theft or be targeted by phishing attacks. Such attacks are a big concern as the researchers indicated that the participants tend to be elderly and therefore not tech-savvy. Moreover, a data leak can decrease confidence in the platform and discourage citizens from participating in research, thwarting scientific progress and the researchers themselves can face sanctions from ethical boards and legal action. Therefore, we have to pay special attention to authentication, authorisation and prevent any vulnerabilities from being introduced.

When it comes to handling personal data, GDPR also comes into scope as a possible legal risk. However, we do believe our system to be compliant with GDPR legislation due to the fact that academic guidelines for the collection, processing and storing of personal data are often stricter than what GDPR requires. In essence, the entire dynamic informed consent procedure guarantees an explicit consent and provides the user with all the tools they need to update, modify or delete their data at any time.

6.2 Operational risks

The most significant concerns relate to the bodily well-being of citizens using the platform. As the exact research that will employ the platform is unknown, it is possible that it will be used to track medicine dosages and how it relates to experienced positive and negative effects. Due to the nature of the application, a hypothetical user might revisit data to decide the dose of the medication that they want to take. However, if there is any corruption of ambiguous display of number (for example, relating to the units or the decimal point) a user might be mislead to take the wrong dose which could cause real-world harm. This is a very serious albeit far-fetched example of a way that the system can cause physical harm to its users. However, we believe that the scenario outlined above is exceedingly unlikely and would prove to be a more reliable solution than other tracking methods. A user might be inclined to track their medication on paper, which would in turn mean that the reliability of the data would depend on many more variables such as organization and handwriting legibility. Therefore, we can confidently conclude that our platform is unlikely to cause any bodily harm to its users and would present an improvement over other possible tracking solutions.

Besides, the citizens must not be exploited for research. While dynamic consent was an issue that was explicitly highlighted by the researchers, special care must be taken to ensure that the users known exactly what their data is being used for. Insufficient measures to ensure the citizens know what they are signing up to do can be heavily scrutinized by regulatory bodies and the general public. Therefore, special care must be taken that the citizens are not only informed but understand what they are consenting to.

6.3 Planning risks

Delivering a working platform on time is an explicit goal of the project. Obvious concerns around planning automatically arise:

- Do we have enough people to build the system?
- Are we able to finish a product on time?

- How do we deal with uncertainties like sickness or unavailability of a team member?
- Is the client and are supervisors available throughout the project

These risks are managed by first concluding we do have enough people to build the envisioned system. If we will be able to deliver a product depends in large part on if the requirements are well-defined beforehand. Scope creep [3] is a natural risk that can be mitigated by sticking to (just) fulfilling the requirements set out in this proposal. Absence of team members cannot be predicted, but can be dealt with accordingly when the time comes, for example by weekly evaluating where we stand. Upon first contact with the clients and supervisor, we investigated their availability throughout the project to avoid any surprises. Taking all of the former into account, we believe there is no reason it is not to be expected we will deliver a working product on time.

References

[1] Maria LaVictoire and Nick Everhart. "A touch screen button size and spacing study with older adults". In: *SpringerLink* (Jan. 1970). URL: https://link.springer.com/chapter/10.1007/978-3-642-02707-9_29.

[2] E.F. Loos et al. "User-friendly websites in the eyes of young and old people". In: *Proceedings of A transdisciplinary conference organised by COST Action 298 "Participation in the Broadband Society", Copenhagen, Denmark, 13th-15th May 2009* (May 2009). URL: https://dspace.library.uu.nl/handle/1874/40464.

[3] Sergiyo Sayago and J. Blat. "About the relevance of accessibility barriers in the everyday interactions of older people with the web: Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)". In: *ACM Other conferences* (Apr. 2009). URL: https://dl.acm.org/doi/abs/10.1145/1535654.1535682.

[4] Shari Trewin. "Physical impairment". In: *SpringerLink* (Jan. 1970). URL: https://link.springer.com/chapter/10.1007/978-1-84800-050-6_4.

[5] Panayiotis Zaphiris, Sri Kurniawan, and Mariya Ghiawadwala. "A systematic approach to the development of research-based web design guidelines for older people - universal access in the information society". In: *SpringerLink* (Nov. 2006). URL: https://link.springer.com/article/10.1007/s10209-006-0054-8.

Appendices

A Requirements

The following section contains all the requirements as specified by the client, classified according to the MoSCoW method. Each requirement can be tagged as Must be implemented (M), Should be implemented (S), Could be implemented (C) or Won't be implemented (W).

A.1 Functional Requirements

Functional requirements are concrete, objective requirements the system should satisfy. Most requirements encompass a single feature.

[3] Subtle changes (often many smaller, incremental updates) to requirements that widen the scope of the to-be-built system, sometimes causing the workload to exceed development capacity

A.1.1 Research set-up

a. As a **researcher**, I want to be able to...

1. create my own survey. (M)
2. create a text-box question in a survey. (M)
3. delete questions in a survey. (M)
4. edit questions in a survey. (M)
5. add a title and description to a survey. (M)
6. set a period in which the survey is available. (M)
7. set an interval after which a response is desired, so that participants can fill out the same survey every day, week or any other period. (M)
8. read information about how to use the platform. (M)
9. specify a date at which the data will be extracted for analysis so that participants know when they lose the right to revoke consent. (M)
10. require the use of 2FA. (M)
11. create a number question in a survey. (M)
12. create a short text question in a survey. (S)
13. create a slider question in a survey. (S)
14. create a checkbox question in a survey. (S)
15. create a drop-down question in a survey. (S)
16. create a multiple choice question in a survey. (S)
17. set whether a survey is public or private, which determines whether a citizen can find the survey without a direct link. (S)
18. create an image question in a survey, in which respondents can select one or multiple images as answer to a question. (S)
19. create a date question in a survey. (S)
20. create a time question in a survey. (S)
21. add text in-between questions in a survey, so that I can provide additional information or explanation. (S)
22. add images to any question or in-between section, so that I can provide visual context. (S)
23. change the order of questions in a survey. (S)
24. make a question optional to answer or not. (S)
25. set a default answer to a question. (S)
26. add (and remove) another researcher so they can extract data from a survey. (S)
27. add (and remove) another researcher so they can manage a survey. (S)
28. duplicate questions in a survey. (C)
29. create a file-upload question in a survey. (C)

30. create a multiple choice grid question in a survey. (C)
31. create a checkbox grid question in a survey. (C)
32. add videos to any question or in-between section, so that I can provide visual context. (C)
33. duplicate surveys, so that I have a basis to work from if the survey I am planning to make is very similar to another survey. (C)
34. add feedback to an answer, so that when a respondent selects an answer, the corresponding feedback is displayed. (C)
35. add a confirmation message. (C)

A.1.2 Onboarding process

a. As a **citizen**, I want to be able to...

1. sign up for the platform with my email address and a password. (M)
2. for each study I consider participating in, read information about that study. (M)
3. read information about how to use the platform. (M)
4. read information about a survey and how my data will be used. (M)
5. choose whether I give consent for my data to be used in research. (M)
6. read information about dynamic informed consent. (M)
7. easily provide my consent. (M)
8. include consent to be contacted about other similar surveys in my dynamic consent. (S)
9. look for surveys to participate in. (S)
10. sign up for the platform with my social media (Google, Facebook) account. (C)

b. As a **researcher**, I want to be able to...

1. generate a URL to the registration page of a survey. (S)
2. generate a QR code to the registration page of a survey. (C)

A.1.3 Data ingestion

a. As a **citizen**, I want to be able to...

1. fill out a survey. (M)
2. be reminded of a survey whenever it is available to be filled out. (M)
3. see data entries from a specific date/time. (M)
4. log in to my account to view all surveys I currently participate in. (M)
5. see when the deadline for granting or revoking consent is. (M)
6. edit my previous survey submission, so that I can correct errors. (S)
7. fill out a survey without logging in. (S)
8. view accumulated data from a survey in a graph. (S)
9. display multiple data sets in a single graph. (S)

10. export my past data. (S)
11. see a progress bar for the duration of the survey. (C)
12. connect my hospital data to the platform. (W)
13. share my data from health tracking devices with the platform automatically. (W)

b. As a **researcher**, I want to be able to...

1. track progress and status of (groups of) citizens while the survey runs. (M)

A.1.4 Data processing

a. As a **researcher**, I want to be able to...

1. extract data from a survey from a specific period as a CSV file. (M)
2. view simple aggregated statistics from a research. (C)

A.1.5 Feedback and communication with citizens

a. As a **citizen**, I want to be able to...

1. send a message to the researchers from a survey. (M)
2. receive a notification when the deadline for revoking or granting consent passes. (S)
3. request the creation of a new survey. (S)
4. ask questions regarding the platform. (S)

b. As a **researcher**, I want to be able to...

1. respond to messages from citizens. (M)
2. send a message to all participants from a specific study. (M)
3. send a message to all participants from a specific study that agreed to be contacted about new surveys similar to that one. (S)
4. send a message to individual citizens. (S)
5. send a message to a selection of participants form a specific study. (S)
6. (automatically) inform citizens after a survey is completed. (S)

A.1.6 Privacy

a. As a **citizen**, I want to...

1. be able to revoke my consent regarding the usage of my data in a survey. (M)
2. be able to revoke my consent regarding the usage of new data I supply in a survey. (M)
3. be able to delete all my data from a survey. (M)
4. be able to retroactively give consent regarding the usage of data in a survey. (M)
5. be able to give consent regarding the usage of new data I supply in a survey. (M)
6. log in before I can view my past data. (M)
7. be able to choose whether to give consent regarding the usage of my data in other studies besides the one the survey was created for. (S)

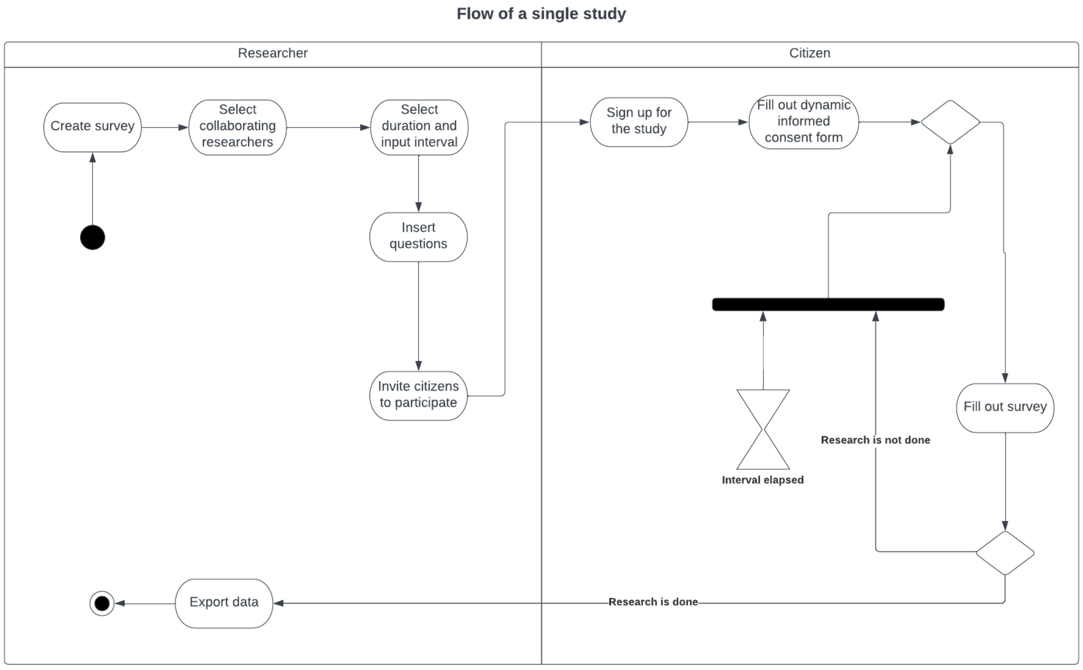
A.2 Non-Functional requirements

Non-functional requirements do not directly affect one feature of the system, but rather represent a cross-cutting concern that must be kept in mind and satisfied throughout the system.

1. Citizens should remain anonymous when contacting researchers.
2. Citizens should remain anonymous when filling out a survey.
3. The system should minimize scrolling for the citizen.
4. The system should display high contrasting colors for the citizen.
5. Screen readers should work on each page.
6. The system must be fully navigable by keyboard.
7. Images should have alternative text.
8. The system must be accessible from mobile phones and/or tablets.
9. Security-related best practices should be followed (such as specified in the OWASP top 10). [4]

B Diagrams

B.1 Activity diagram



Flow of a single study

Figure 3: Activity diagram of the system

[4] https://owasp.org/www-project-top-ten/