Institute for Visualization and Interactive Systems

University of Stuttgart
Universitätsstraße 38
D–70569 Stuttgart

Bachelorarbeit Nr. 320

# Designing for Motivation: Encouraging Volunteers to Participate in Data Collection

Patrick Tobien

| | |
|---|---|
| **Course of Study:** | Informatik |
| **Examiner:** | Prof. Dr. Albrecht Schmidt |
| **Supervisor:** | Dipl.-Inf. Lars Lischke,<br>Marco Hirsch, M.Sc.,<br>Robert Krüger, M.Sc. |
| **Commenced:** | April 20, 2016 |
| **Completed:** | September 20, 2016 |
| **CR-Classification:** | H.5.2 |

# Abstract

Data collection is an important aspect of all research, be it in the fields of sciences, humanities or business. Participatory data collection, in particular, poses a challenge to all fields of study as the required data for modern research often exceeds the amount that is provided by voluntary participants. It is therefore essential to find more people who are willing to submit their data.

This thesis suggests and investigates potential designs for mobile applications that motivate volunteers and provide incentives for participation. First, multiple design methods are introduced and explained ranging from reward-based designs, over gamification and storytelling, to the integration of social networks. Five designs incorporating these methods are then suggested and evaluated via an online survey. The designs focus on a mobile application that uses symptom questionnaires to predict the spreading of influenza-like diseases. Utilizing the results of the survey, a prototype using the design with the highest potential is implemented on Android to further assess its effectiveness. The prototype is evaluated in a user study using ContraVision – interviews about two videos showing the positive and negative aspects of the design. Results of this work provide potential design methods that increase motivation to participate in data collection. The acquisition of more data allows for more research questions to be answered and ensures correct assumptions for further research.

# Kurzfassung

Datensammlung ist ein essenzieller Aspekt in jeglicher Forschung, egal ob im Bereich der Naturwissenschaften, Geisteswissenschaften oder im Geschäftswesen. Insbesondere partizipative Datensammlung stellt eine Herausforderung für alle Forschungsbereiche dar, da die Menge an erforderlichen Daten oftmals die Menge übersteigt, die von freiwilligen Teilnehmern bereitgestellt wird. Es ist darum sehr wichtig, mehr Personen zu finden, die bereit sind, ihre Daten zur Verfügung zu stellen.

Diese Arbeit untersucht und schlägt potenzielle Designs für mobile Anwendungen vor, die Freiwillige motivieren und Anreize zur Teilnahme liefern sollen. Zunächst werden mehrere Design Methoden vorgestellt und erklärt. Diese erstrecken sich von belohnungsbasierten Designs über Gamification und Storytelling bis hin zur Integration von sozialen Netzwerken. Fünf Designs, die diese Methoden umsetzen, werden daraufhin vorgeschlagen und mit Hilfe einer online Umfrage evaluiert. Die Designs beschränken sich auf eine mobile Anwendung, die Umfragen zu Symptomen durchführt, um die Ausbreitung von grippeähnlichen Krankheiten vorherzusagen. Unter Berücksichtigung der Umfrageergebnisse wird anschließend ein Prototyp für Android implementiert, der das Design mit dem höchsten Potenzial umsetzen soll, um die Effektivität dessen zu bestimmen. Der Prototyp wird dann mit einer Benutzerstudie mit Hilfe von ContraVision evaluiert. Dazu werden Interviews durchgeführt, nachdem zwei Videos, die die positiven und negativen Aspekte des Designs darlegen, gezeigt wurden. Die Ergebnisse dieser Arbeit liefern einen Überblick über potenzielle Design-Methoden, die Motivation zur Teilnahme an Datensammlung bieten. Die Aneignung von mehr Daten ermöglichen die Beantwortung von mehr relevanten Forschungsfragen und gewährleistet korrekte Annahmen für die anschließende Weiterforschung.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

In any given area of interest, in order to find answers to relevant questions or make predictions about future probabilities or trends, it is necessary to gather sufficient information. Said information is integral in research and imperative in order to make informed business decisions. For instance, information might be retrieved from mobile applications, websites and surveys in order to evaluate and improve retail sales.

The Oxford English Dictionary defines data collection as "*the activity of collecting information that can be used to find out about a particular subject*". Thus, it describes the systematic approach of gathering and measuring information. Accurate data collection is important as well as detrimental in many disciplines including sciences, humanities and business. Typical methods of data collection consist of observations, surveys, interviews and focus groups. Additionally, today's collection methods include the use of mobile devices, website traffic, internet activity and other technological sources. Mobile devices also offer a variety of sensors and measurements that provide additional relevant data, depending on the research field.

This study focuses on participatory data collection, i.e. on gathering data from people via active contribution. More specifically, this thesis intents to find ways to motivate people to contribute their data via mobile applications. The central question of this research is whether people can be motivated to participate in data collection via design incentives. The purpose of this study is to find possible design methods that could improve participation and to evaluate them.

Even though there are many existing collection methods that are often simple (e.g. mobile or web surveys), attaining the necessary data is not always as easy as it looks. More often than not, people are not interested in providing personal information, especially if it involves personal effort with no immediate gain in return. Additionally, gathering information of people (anonymous or not) also always invades their privacy. People rarely want to be or feel observed, analyzed or judged and usually care about their privacy. Disregarding people's privacy is also against the law in most countries.

Therefore, it is important to find enough people to gather said data from. And it is just as important to do so without invading people's privacy, i.e. by finding volunteers who are willing to submit the necessary data freely and while being informed.

Since the subject of gathering health data is an important and non-redundant aspect in past and modern society, it was chosen as a thematic platform for this thesis. In particular, the prediction of flu like illnesses (ILI) constitutes a significant research area. Therefore, this study looks at participatory approaches for health data collection and possible improvements. Prediction of disease spreading in general is the first, and possibly most important, step in finding ways to alleviate their effects [Lon+05] and was chosen because it is a perpetually important research area that can be improved with enough data. The topic also keeps appearing as the spreading of diseases make the news (avian flu, swine flu, Ebola, BSE, etc.).

A simple attempt to acquire more data for said prediction is to ask people about their health (i.e. whether they show any symptoms and if yes, which) on a regular basis. Above all, this thesis focuses on collection methods via regular surveys on a mobile platform, whereby the collected data is used to predict influenza spreading. A mobile application is easy to distribute, easy to use and has the ability to reach many possible participants. The downside is that few people are interested in reporting their symptoms on a regular basis, especially if they feel healthy – even if the task itself would be simple to do.

Typically, people are rewarded for their participation in data collection with monetary compensations. Some research was done on methodologies that increase response rates in, for instance, surveys targeted at healthcare professionals [VJW07]. However, those were generally focused on monetary rewards as well as on the medium, i.e. e-mail, postal, fax, telephone, live interviews, etc. This study focuses on the design instead, i.e. on techniques that can be applied to any device that is used for participatory data collection. In particular, these techniques are used to improve the prediction of flu spreading, which is an important research area because influenza is a pandemic with yearly outbrakes resulting in about 250,000 to 500,000 deaths [Org+14]. In theory, flu shots can be manufactured and distributed, but they differ between different influenza types and have to be available beforehand. Unfortunately, it is difficult to predict which type of influenza will spread in a given season. Studies have shown that early detection enables preventive measures that can contain epidemics [Fer+05], [Lon+05].

The findings of this thesis play an important role in all research as data is necessary to generate knowledge and one of the major methods to acquire said data is to gather it from people. The more data researchers have the more relevant research questions can be answered. Correct and sufficient data ensures that researchers make correct assumptions for further studies. This work provides an evaluation of as well as suggestions for potential design methods that increase participation rates for data collection. The results of this thesis could be applied to numerous future data collection processes that require active participation. This work thereby contributes to all fields of study as a starting point to improve existing data gathering methods.

## 1.1 Organization of this Thesis

Following this introduction about data collection and health data in general, a short overview of background information and related work is presented. This includes preceding work on data collection techniques and issues, and on existing methods that use participatory approaches such as surveys and alike to improve disease spreading prediction, in particular that of influenza. Additionally, a short introduction of motivation from a psychological standpoint is given in order to better understand what drives people to do something or to act upon.

Following that, five different design techniques are introduced and explained as possible incentives to participate in data collection. The techniques are applied to a theoretical mobile application that utilizes surveys in order to predict influenza spreading. Mock-ups are designed in order to learn what might work as a motivating incentive and what might not. The resulting prototypes are evaluated via an online survey and compared to a prototype that does not use any of the suggested designs. Results are compared and assessed with respect to possible issues.

Subsequently, the most promising looking mock-up is implemented as a usable prototype on Android. Said prototype is then evaluated using ContraVision [Man+10]. Based on these results, a quick outlook on this study's relevance for future work is presented.

# 2 Background and Related Work

As mentioned above, surveys are an important tool in data collection. Previous studies have found reasons why survey participation is poor in many cases. Surveys of physicians or nurses, for instance, are an important aspect in healthcare research. Since participation rates are generally low, four main reasons have been indentified why healthcare professionals are reluctant to respond to surveys: “time constraints, poor received value of the survey, issues of confidentiality, and perception bias” [Sud85]. Methodologies for improving response rates were discussed in Vangeest et al. [VJW07]. Monetary rewards, such as even small financial incentives, were found to be effective. From a design perspective, postal and telephone strategies proved to be more effective than fax or web-based approaches, as did questionnaires that were designed to be brief and personalized. A comprehensive comparison on how to motivate nurses to participate in surveys, whether by incentives or design approaches, was given in Vangeest et al. [VJ11] with similar results overall: Financial incentives prove to be effective. Phone and mail based surveys give better participation results than web-based approaches.

Since this thesis focuses on motivating people to participate in (active) mobile collection of ILI data, a quick characterization of motivation as a psychological construct is given in the next section. A subsequent introduction about crowdsourcing is given, followed by a brief overview of existing mobile applications that collect data. Lastly, a summary of existing approaches to predict disease spreading, in particular that of influenza, is shown.

## 2.1 Motivation as a Psychological Construct

To assess how motivation can be triggered in people we first have to understand what motivation is. There are various psychological theories and models that describe motivation in human behaviour. A brief overview of the most popular models is given in the following.

A very simple model is provided by the distinction of intrinsic and extrinsic motivation [RD00]. Intrinsic motivation describes inner drives and desires that compel a person to seek new challenges or knowledge. The motivation comes from a feeling of joy from the activity itself not because of a feeling of obligation or the desire to achieve a specific goal. In contrast to that, extrinsic motivation refers to a desire to produce a certain outcome. Outer influences generate extrinsic motivation, such as rewards (grades, money, etc.) for achievement or punishment for failure. The prospect of recognition or popularity are also considered extrinsic incentives.

Behaviorists such as Skinner and Pavlov focus on a systematic approach on human behaviour, namely the principle of stimulus and reaction. Behaviorists believe that a person's behaviour is the result of their environment (i.e. outer factors), past or present. A very widespread behavioral model justifies certain behavior via conditioning [PA03]. As such, an individual might associate emotions such as pain or pleasure to certain actions or their outcomes. In that sense, motivation is caused by previous (similar) experiences and the resulting favorable stimulus (the removal of pain is a favorable stimulus). Similarly, behaviorist theories also cover incentives (see above), as incentives such as rewards are basically stimuli either from oneself or from other people. The distinction between intrinsic and extrinsic is hereby irrelevant. Immediate rewards seem to provide a bigger incentive than rewards further in the future.

The typical economics view of motivation consists of people as rational agents that try to maximise their payoff whereby the world consists of several positive or negative payoffs. In order to make people behave in a certain way (extrinsic) incentives have to be employed. This so called "mechanism design" [Var95] is also a very active field in game theory. A possible method would be to reward users with a return service, most simply (virtual) financial rewards for their participation. An example can be found in Yahoo! Answers or Google Answers. The latter operated between 2001 and 2006 and used real money as reward for helpful contributions to the community. However, the system was riddled with people trying to "game" the system which lowered the quality of responses and, following a lot of user's complaints, ultimately led to Google Answer's demise. Yahoo! Answers, in an attempt to avoid the problems that Google Answers had to face, uses a rewards system based on visible badges and statuses that represent a user's amount and quality of contribution, such as "power user" or "top contributor".

It is important to keep the various theories in mind when dealing with the generation of motivation for volunteers in data collection. However, since it is impossible or at least unfeasible to meet an individual's personal requirements to provide motivation to participate in data collection, a simple approach with an incentive based theory with the distinction between intrinsic and extrinsic motivation, whereby we will use a behavioral outlook instead of a cognitive one, will be sufficient.

Ideally, we want to combine various motives – or in our case, various incentives. Extrinsic motivation is, in theory, easy to achieve, e.g. via rewards for certain goals or possibly even punishment for non-fulfillment. Intrinsic motivation is evidently more difficult to evoke.

## 2.2 Crowdsourcing

Crowdsourcing is a means by which problems are solved, content is created or services are provided by a large group of people, usually an online community. It constitutes an effective, cheap and very popular alternative to outsourcing [How06]. In the context of this study, crowdsourcing is used to obtain the necessary ILI data in order to predict flu epidemics.

One of most prominent examples of crowdsourcing is Amazon Mechanical Turk, an Internet marketplace that enables so-called Requesters to employ human intelligence for tasks that can not be performed (cost-effectively) by computers. This proves to be an inexpensive and fast method of obtaining high-quality data [BKG11]. The financial compensation received for the tasks is usually low, since the tasks are small and increased payments do not increase the quality of the data [MW10]. Another popular example is the use of crowdfunding, i.e. raising vast monetary contributions from a large group of people [Mol14]. The most prominent crowdfunding site is Kickstarter with over 2.5 billion $ raised in pledges and over 110,000 projects funded successfully [Kic16].

Crowdsourcing as a data collection method proves to be important and effective. For instance, with the help of Social Media, crowdsourcing applications can be applied to disaster relief coordination [Gao+11]. Its application in market-related activities ranges from employing individuals to complete menial repetitive tasks for small financial compensation to finding solutions from a crowd of experts [Whi09].

Goncalves et al. [Gon+13] examined crowdsourcing on public displays. In their study they found that public displays are suitable for altruistic use. They adjusted the displayed text according to motivational constructs such as “community” (help others by performing the given task) and “enjoyment” (accomplish X by performing the given task). They found that careful design of those incentives can be benefitial for participation.

On mobile devices, crowdsourcing can either be done participatory, i.e. with user generated data, or opportunistic, i.e. from sensors and automatically performed computations. Smartphones prove to be an ideal platform for crowdsourcing [Cha+12]. Reward-based incentives for participation in crowdsourcing applications on smartphones have been examined by Yang et al. [Yan+12].

## 2.3 Mobile Applications

As mentioned before, this study suggests a mobile application to gather ILI data from a large userbase via crowdsourcing. The idea of using mobile devices to collect data is anything but new. Today's mobile phones are capable of capturing, processing and transmitting location, image, acoustic, motion and other data. Methods or applications to collect data have been proposed and developed numerous times [Bur+06], [JBE07], [Sin07], and privacy issues have been discussed [Shi09], [BSM12]. Various applications use sensory data from mobile devices in conjunction with participatory systems in order to collect real world data. GIS Cloud's Mobile Data Collection, for instance, enables users to collect real-time field location, multimedia and attribute data [GIS16].

Various mobile applications reward users for letting them collect passive data such as accelerometer, GPS, app usage, battery life, operation system, gyroscope and other data. Panel App [Pla16] runs in the background and rewards points which can be exchanged for gift cards or other prizes. Cross Media Panel [Inc16b] is a market research approach on behalf of Google that tracks user's web behaviour and app usage in order to improve products like Chrome, Android and YouTube.

Mobile applications are also used to conduct surveys. With SurveySwipe [Ana16] people can distribute and complete surveys for small rewards in addition to signing up for voluntary passive data collection. Typical application areas are surveys among employees to improve business decisions and among consumers to improve sales. Usually there are no direct incentives to participate other than it being a compulsory deed. Sometimes, however, monetary rewards are offered as compensation. Google Opinion Reward, for instance, offers Play credits in return for answering few questions around once a week, which can be spent freely on the Google Play Store [Inc16e].

From an optical design perspective, Kukka et al. [Kuk+13] have examined what makes people use public displays. They reported that text is more effective in incentivizing interaction than icons. Color is more effective than greyscale and static signals are better than animated.

In order to offer incentives that improve participation in data collection, an engaging user experience has to be designed. The increasing importance of user experience has been discussed by Hassenzahl and Tractinsky [HT06].

## 2.4 Prediction of Disease Spreading

Prediction of disease spreading is perfectly suited for an application of possible incentives for participatory data collection, since it can be easily improved if enough people are willing to submit the necessary data on a regular basis.

Traditionally, influenza spreading is predicted by the amount of people visiting the doctor for flu shots in a specific area and by relying on previous experiences. Additionally, some innovative systems have been proposed like evaluating telephone triage data [EHW03] or using over-the-counter pharmaceutical sales [Mag03].

However, with the rise of the internet and mobile platforms, additional methods have emerged. Since 2008, for instance, search terms are used by Google in order to estimate the spread of influenza and dengue fever and the resulting data was published until late 2015 [Inc15]. These Google Flu Trends have been shown to provide substantial improvements over traditional methods [Dug+13].

Another popular method is the use of social media, mostly Twitter, as a data source for disease spreading. Various approaches have been made by different research teams to use said data for flu prediction [AMM11; Ach+11]. For that, the Twitter datastream is analyzed and processed: Flu-related tweets that are also geotagged are filtered in order to create a map that shows the distribution of these tweets and how they vary over time. Li and Cardie [LC13] also modeled the dynamics of the disease into four phases (non-epidemic, rising, stationary, decreasing) in order to spot a switch from one phase to another as early as possible.

Multiple networks use influenza-like illness (ILI) self-reporting systems and thereby employ participatory data to improve prediction of spreading: Flu Near You (FNY) is an existing approach in the United States and stands out as it offers a fairly popular mobile application [Chu+12a]. Users are asked to submit their symptoms (if they have any) on a weekly basis and FNY uses the gathered data for predictions. In August 2012 over 9300 people across the United States participated. According to their website, the number of participants increased to over 60,000 in 2016 [Har16].

The in 2008 established web-based network Influenzanet also depends on the participation of the population in order to produce results that are better than those of state-of-the-art influenza surveillance systems [Pao+14]. Influenzanet is a European network and provides real-time information on ILI cases throughout the continent. Data is gathered via weekly online surveys and depends on participation in large numbers – similar to Flu Near You.

Participatory systems have previously been shown to be effective, such as in the case of malaria surveillance in urban India 2010-2011, where people were compensated via micro-monetary rewards [Chu+12b]. The following chapters thus focus on improving such participatory approaches by designing a mobile application that offers motivational incentives other than payoff.

## 2.5 Submitted Poster Based on this Work

Based on the findings of this study, a poster has been submitted to the *2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp 2016):

> Tobien, P., Lischke, L., Hirsch, M., Krüger, R., Lukowicz P., Schmidt, A.:
> Engaging People to Participate in Data Collection
> *2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp 2016)*

# 3 Design

In order to find incentives for a mobile application that motivate people to participate in data collection, various techniques are researched and then designed as different prototypes to compare to each other. As mentioned in the introduction, the basic design idea is an application that uses weekly surveys to gather data for ILI spreading. This includes a report function, a profile with general user information and an overview of existing reports (a map or similar). In the following section, various techniques are introduced that can be applied to a mobile application with hopes of improving the response rate of users.

## 3.1 Techniques

Four different design techniques for mobile applications to engage volunteers in data collection are proposed in this study. Mockups for each technique are designed as prototypes and later evaluated against a clean application with none of the techniques applied to it. The results are used to find the best approach and continue with an implemented prototype to further assess its effectiveness.

### 3.1.1 Gamification

The first technique and best covered approach in previous research is gamification. Gamification is the application of game principles and mechanics in design fields that are not games in order to improve user motivation and engagement [Det+11].

There is a lot of research on gamification and it has been regularly applied as a design technique in order to improve user engagement, perseverance and loyalty. Typically, gamification involves design elements such as badges or achievements (rewards), experience or levels (status) and comparison with other users via leaderboards or similar (competition). These are easy to implement and provide an excellent source of extrinsic motivation.

Gamification has already been shown to be effective for data collection. It proved especially successful for tasks that require human intelligence if combined with crowd-sourcing. Flatla et al. used game elements to turn calibrating tasks into enjoyable activities [Fla+11]. Three example games were implemented for various calibration tasks. For instance, one of those games used a space-invaders-like mechanic with targets that vary in visibility in order to test a user's ability to see different colors. The gamified calibration tasks proved to be a lot more enjoyable than traditional methods and were strongly favored in their evaluations. Similarly, von Ahn and Dabbish developed a computer game to crowdsource labeling of images [VAD04]. They learned that turning a normally tedious task into something that people actually enjoy far outperforms the development of complicated algorithms for the same tasks. They showed that letting 5000 people play the game 24 hours a day would allow them to label all images indexed by Google (in 2004) in a matter of weeks. Dergousoff and Mandryk [DM15] also combined gamification and crowdsourcing to gather data on human behaviour with a mobile application with mixed results. Users who used the in-game rewards were motivated to do experimental tasks while users who were asked to do cognitive tasks performed worse than on a non-gamified laboratory version.

Gamification has been proven to be not only an effective but also a highly valuable design technique various times. In 2008, a puzzle game called FoldIt was released by the Center for Game Science at the University of Washington. People try to fold proteins as well as possible with selected tools provided by the game. In 2011, after only 10 days, players deciphered the structure of the Mason-Pfizer monkey virus (M-PMV), a virus that causes AIDS [Kha+11]. The problem had been unsolved for 15 years.

There are various known problems with gamification, however, such as users trying to "game" the system, i.e. to maximize their payoff while minimizing their effort which usually leads to a heavy loss of data quality. This is also what happened with Google Answers where users stopped caring about actually helping others but instead tried to maximize their reward.

Modern gamification therefore often tries to incorporate other methods than the usual reward and progression based system. Ideally, intrinsic instead of extrinsic motivation would be generated. It is, however, a lot more difficult to produce with gamification. Gamification also mostly aims at younger demographics. A lot of the older generations can not relate to games and don't see virtual rewards, status or competition as incentives. Since this study aims to collect data related to flu or similar disease spreading, we need to clarify whether we can reach out to the right demographic with this method.

Daniel Cook gave a speech in Octobre 2007 where he made the claim that "if an activity can be learned, a player's performance can be measured and the player can be rewarded or punished in a timely fashion, then any activity that meets these criteria can be turned into a game". As a way to backup his claim, in 2010, he turned Microsoft office into a game. Ribbon-Hero is a game that teaches skills in office products by engaging the user in a game [Coo10].

However, for data collection there is no real skill that has to be acquired apart from checking the right boxes. Nor can the user's performance be measured apart from taking part in the activity (i.e. reporting symptoms) itself. The user can, however, be rewarded for their participation or completion of other tasks such as adding friends. In that sense, turning the application into a real game experience may be impossible, but certain game elements can be borrowed to engage users.

### 3.1.2 Storytelling

Storytelling attempts to improve user engagement by providing a (personalized) story. Entertainment media such as movies, novels and games usually involve some kind of story. A typical story entails specific elements like character, plot, setting, theme and style [Swe15].

Stickdorn and Schneider [Sti+11] claimed that storytelling serves the purpose of sharing insight. However, according to Pruitt and Adlin [PA10] "All Human Beings have an innate need to hear and tell stories and to have a story to live by". Storytelling is an important part of every culture and has many applications such as for political practice or therapeutic effects [Law01]. It has also been researched in the context of user experience by Quesenbery and Brooks [QB10].

Although the incentive to start a book or watch a movie can be extrinsic (e.g. through the recommendation of a friend or a review of a critic), the actual task of doing it usually ends up being intrinsically motivated (i.e. it gives the person a feeling of joy or they have an inherent interest in the story). Of course, there are other elements that engage people in a movie or a book, but more often than not a story plays an important role

With regards to user experience and as a design method these elements would have to be slightly remodeled. Characters could be given by the user or a group of users, the setting would depend on where the user resides and the other elements would depend on what the user wants to tell. For instance, with Snapchat Story the user creates their own story based on pictures (snapshots) and small captions. Those pictures get deleted after 24 hours. The idea is to create a one-day story based on the most interesting events and venues the user went to that day.

MyFitnessPal is a diet and exercise tracking application that determines optimal caloric intake. In addition to that, a diary gets created based on a user's activities that day, such as running for 10 kilometers or eating certain foods. This also creates a story just like a facebook timeline creates a personal story based on where the user went or what they did.

### 3.1.3 Social Networking

Social networks are an important aspect of modern society. 2.34 billion people worldwide use social networks ( 31% of the world population) and 78% of the American population have at least one social network profile. Facebook and Twitter are the most popular representatives of social networks with Facebook alone having about 1.71 billion active users worldwide [Sta16].

Social network sites (SNSs) provide users with a number of benefits such as social and emotional support, the ability to gather information, bond with other people and especially to build social capital [ESL06]. Lampe et al. [LES06] defined two different reasons for the use of SNSs like Facebook. Social searching is the use of SNSs to find informatiom about offline contacts while social browsing describes the use of SNSs to find new online contacts. The former also entails a surveillance function: The ability to "track the actions, beliefs and interests of the larger groups to which they belong" [LES06]. Social networks are also used to find experts and engage in commercial transactions when needed [Kle08].

Social networks are also the focus of many ongoing behavioral or market research studies. Bollen et al. [BMZ11] report that the public mood state can be derived from Twitter feeds and that mood changes can even be used to predict the stock market. Apart from consumer and market research, Twitter has also been used to analyze political debates and to predict election results [Tum+10].

Studies show that using SNSs can stimulate the brain's reward center and even evoke what is called *flow state*, a feeling of full involvement, focus and absolute joy [Mau+11]. In order to exploit these effects, numerous mobile applications integrate social networks by letting users share their achievements, interests or activities. The New York Times conducted a study on why we share content on social media [Bre12]. They identified five major motivators for sharing. 49% of respondents said they want to bring valuable and entertaining content to one another, 68% use it to define themselves, 78% share content because they want to stay connected to people and nourish their relationships, 69% do it for self-fulfillment and 84% do it to support causes or issues they care about.

As a design technique, integrating social networks can help connect people who have an interest in the topic for which data is collected and let them share it with others for the above reasons.

### 3.1.4 Rewards

Rewards have been proven successful multiple times as seen in section 2. Especially financial incentives have been shown to be effective in improving response rates [AN95], [JO98]. However, this study focuses on design incentives which includes only non-monetary rewards. In a broader sense and in the context of a mobile application, rewards can also mean additional information, additional content or access to additional functions and parts of the application.

In order to be effective and meaningful in the context of influenza spreading, the rewards should be health specific. This means that the suggested design ideas are not as general as the above methods but rather fairly specific to the problem at hand. In that sense, two additional versions are designed. One serves as a diagnosis tool that offers a list of possible infections for the reported symptoms. The other uses activity data from another application or from a tracker in order to show possible correlations between one's fitness (activity) and health.

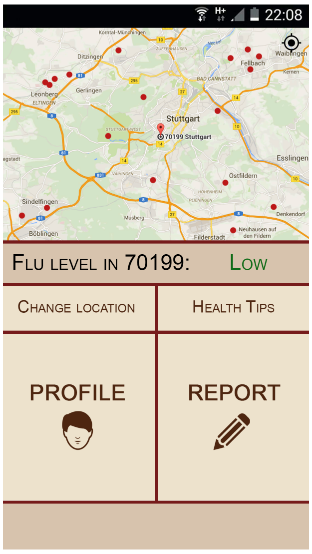
Rewarding concepts appear to be similar to a simple gamified approach, but differ from it in so far that they do not award virtual points or levels but information and statistics with real tangible value for the user. Additionally, the motivation for a reward based approach is solely extrinsic.

It also stands to reason whether these approaches can indeed be called rewards. On one hand, they offer additional information if the user participates in the surveys, i.e. in the data collection: either information about diseases and possible infections or information about one's condition and potential reasons for one's current health. On the other hand, they can be seen as the true purpose of the application, whereby the task of reporting symptoms is merely the means to an end.

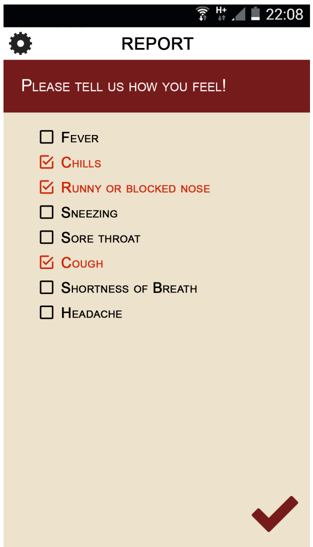
Nevertheless, regardless of the user's intentions, as long as the application is used correctly and regularly, the proposed designs serve as a valuable incentive for participation and are therefore worth evaluating.

## 3.2 Concepts

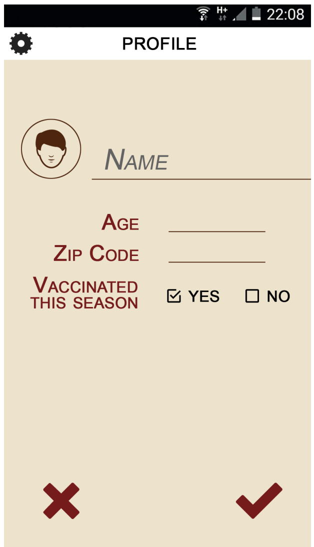
For each of those techniques a concept is suggested that focuses on collecting health specific data via a mobile application. Each application concept is represented by multiple mockups. The basic functionalities are the same in each of those concepts, i.e. the application provides a platform to participate in a weekly survey about the user's symptoms. In case of an infection (of any kind), each symptom can be checked in a list of possible symptoms (see Figure 3.1b). Otherwise, a simple answer stating the user




(a) Main screen with heat map and navigation.

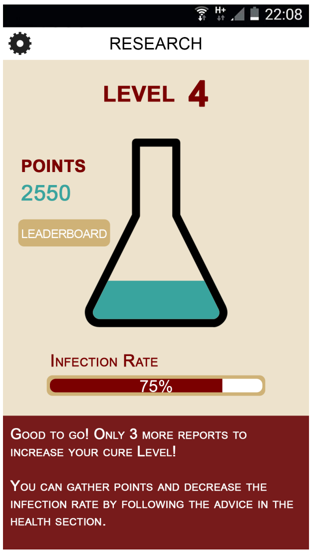



(b) Survey screen with examples of symptoms to check.



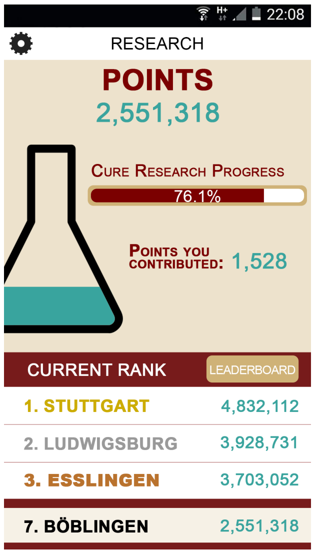

(c) User profile with username and general information.

**Figure 3.1:** Mockups of the standard application with no design techniques applied yet.



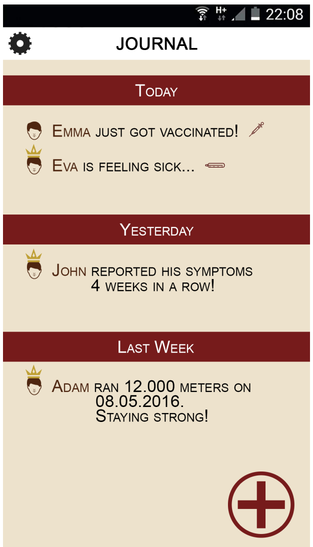

(a) Research screen that shows the current vaccine level and gathered points as well as leaderboards.


(b) Profile with gathered badges.

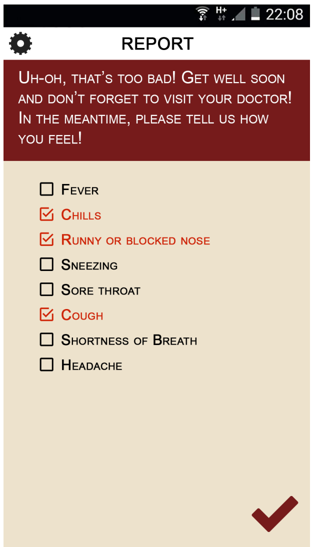



(c) Research teams by area compete against each other.

**Figure 3.2:** Mockups of the gamified application. Points and levels are gained through participation in the weekly survey and completing sidequests.




**(a)** Short journal entries of friends that include valuable information on the user's health.

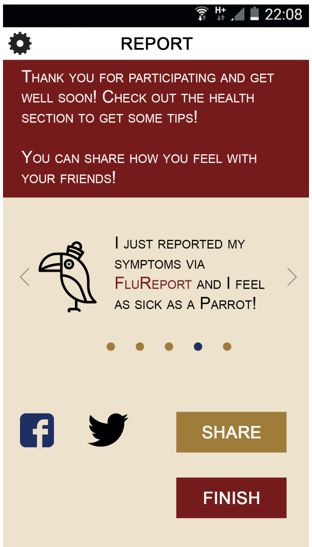



**(b)** Report screen with textual support. The info window speaks to the user and expresses sympathy while humbly asking for the user's symptoms.

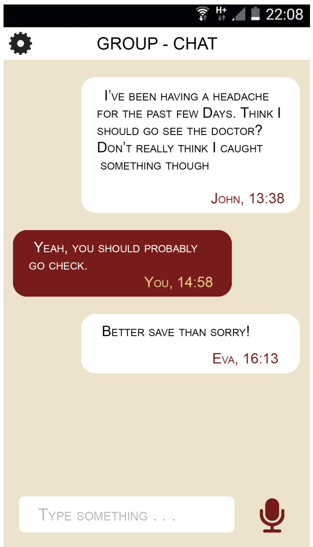
**Figure 3.3:** Mockup of a storytelling design with a journal and textual elements that include the user as a character.

is healthy suffices. Additionally, there is a heatmap on the main screen which shows reports in one's immediate vicinity (see Figure 3.1a) and a profile for some general user information such as sex and age (see Figure 3.1c). The aforementioned functionalities comprise a basic influenza report application.

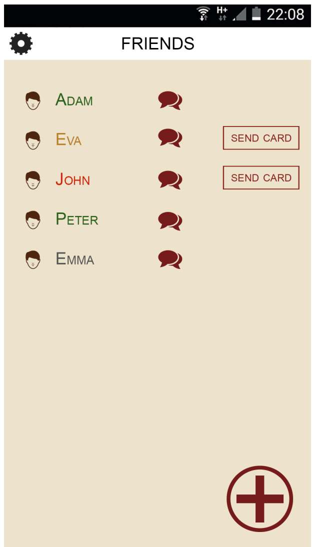
Figure 3.2 shows the gamified approach. It engages the user in a virtual vaccine research, where they have to find a cure to a fictional disease as a part of a research team (see Figure 3.2a). In addition to the basic functions of the application, the user gains points for participating in the weekly survey. For every three consecutive weeks in which the user reported their symptoms, the vaccine improves to a new research stage. Additionally, badges and points are awarded for completing certain quests such as filling out profile information, following health tips and participating a certain number of times in the survey (see Figure 3.2b). The gathered points and research stages are shown in a leaderboard to invoke competition. Another variant incorporates multiplayer aspects by assigning users to a research team by area. Each area then competes against other areas by adding up the individually gathered points and research levels (see Figure 3.2c).




(a) Screen that follows the weekly report. Options to share how the user is feeling on social networks are provided.




(b) Chat room in order to exchange information with friends or family and receive help with health specific topics.



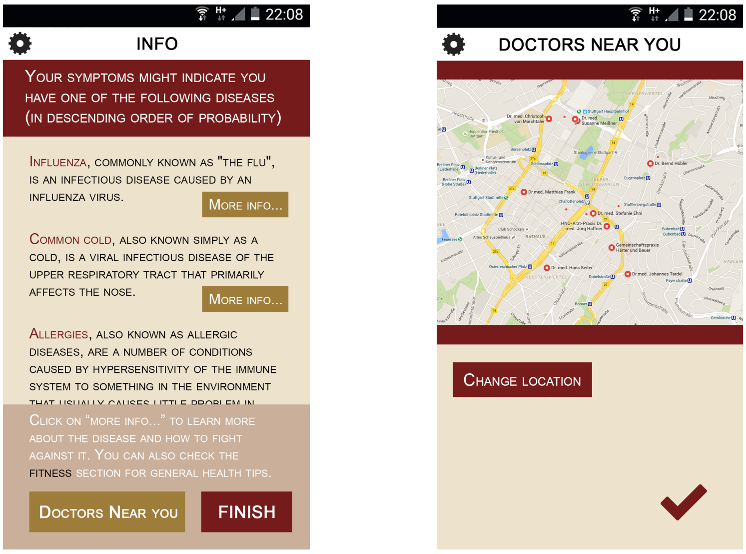

(c) Friendlist with the ability to chat with others. The user can also see who reported sick, who might have a flu and who did not report at all.

**Figure 3.4:** Integration of social networks with the ability to create one's own network vie chatrooms and friendlists.

This also means that friends likely play together in the same research team which might provide another incentive to be invested in the game and application.

The storytelling variant of the application provides the user with the ability to add entries in a journal. Users can tell their friends how they feel, whether they are sick or what activities they did recently (see Figure 3.3a). Similar to a Facebook timeline, this creates a user-specific story about relevant health issues and information. Additionally, task descriptions and menus are stylistically rewritten in a way that engages the user, e.g. by addressing the user personally, congratulating them or expressing sympathy with them (see Figure 3.3b). The received messages change over time, depending on the user's responses. The user is therefore included as a character in the application's "story" which bonds the user emotionally in addition to providing personalized information and feedback.

Integration of SNSs is realized by enabling the user to share their symptoms and statuses on social networks such as facebook and twitter. The user can also talk to other users about topics that relate to the application and health in general, which might enable the



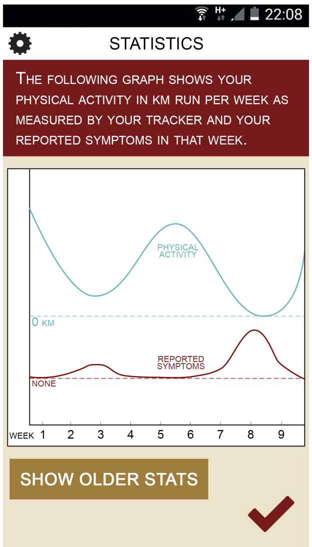

**(a)** A list of possible illnesses according to the reported symptoms in descending order of probability.

**(b)** A map with known physicians near the user.

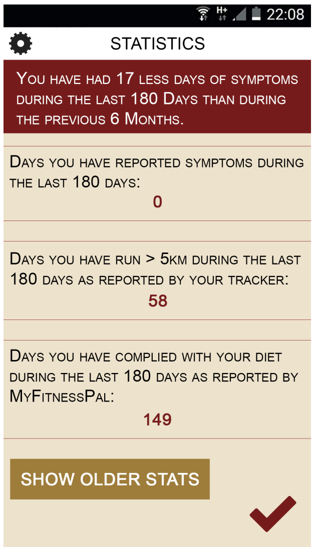
**Figure 3.5:** Mockup of a diagnosis tool that suggests possible infections to the given symptons as well as physicians in the near vicinity. The definitions and information are pulled from wikipedia.

userbase to form a helpful community in relation to these topics (or others). Therefore, the user can share their usage of the application, their symptoms as well as their general condition in order to inform their peers, gain sympathy or even receive help from friends and family (see Figure 3.4a). An internal friendlist and individual chat rooms (see Figure 3.4b) further create possibilities to share and communicate with people of their choosing. If desirable, friends can see the user's status and health conditions in the friendlist (see Figure 3.4c).

As mentioned before, the versions that reward users do not consist of financial compensations. Instead, valuable information within the application is provided. The first approach offers a list of possible diseases for the reported symptoms, using wikipedia articles and sorting them by probability (see Figure 3.5a). The payoff for participating in the weekly survey is therefore a (possible) diagnosis in addition to information and tips on how to recover. Additionally, there is an option to show physicians near the user's location (see Figure 3.5b). The tool is proposed as an additional personalized and informative source of health related topics. It is not supposed to replace traditional health checks.




(a) A graph that shows the possible correlation between physical activity and health.




(b) Some additional textual statistics in order to get an idea about one's progress or trend.

**Figure 3.6:** A mockup of a design that uses otherwise gathered activity data to find correlations between the user's fitness and health.

The second suggestion gathers data from an additional activity tracker to then create diagrams (see Figure 3.6a) and textual statistics (see Figure 3.6b) for a specific amount of time which lets the user see correlations (or not) between their activities or diet and their general health. Thus, the reward consists of information about the user's fitness and health progression and their relation. This also serves an educational process, where the user can learn about their personal patterns that do or do not lead to an infection.

## 3.3 Online Survey

An elaborate online survey has been conducted via Google Forms [Inc16d] over a time period of 4 weeks to evaluate the suggested design techniques and determine the general willingness to participate in data collection.

Initially, a few questions about smartphone and mobile application usage and concerns were asked. Subsequently, each design technique was introduced with a few mockups and a short description. A five-point Likert scale was then used for statements with which the respondents could either agree or disagree. For comparison purposes a basic application without any of the aforementioned design techniques was also shown.

Every proposal is followed by a few design-specific statements and some general statements. Participants also have to give a rating on how much they think it would improve the frequency of them reporting their symptoms.

### 3.3.1 Results

51 people participated in the online survey. The respondents were aged between 15 and 77 with a mean of 27 years of age and a standard deviation of 10.44. 45.1% of respondents were female, while 52.9% were male (2.0% reported other). 51.0% of respondents have a Bachelors degree or higher (Master's or PhD degree) which means the survey was slightly biased towards responses with a background of higher education. All respondents reported that they use a smartphone at least on a daily basis and have used at least a moderate amount of mobile applications before.

Based on a Likert scale of one (strongly disagree) to five (strongly agree), the respondents overwhelmingly agreed that health and wellbeing is an important topic to them with a mean score of 4.12 on the Likert scale. Likewise, with a mean of 3.20 respondents also reported that disease spreading, in particular influenza spreading, is an important topic to them. Considering this, it is not surprising that 56.9% of respondents claimed that the contribution to public good is enough incentive for them to participate in data collection via a mobile application. However, 51.0% of respondents also reported that monetary compensation would be a good incentive for them to participate and 43.1% said they would be motivated through access to additional services or information as a reward of participation.

Following that, the survey results show that at least some of the suggested design techniques could improve voluntary participation in data collection or, more specifically, motivate users to report their symptoms on a weekly basis. 53.0% of respondents (strongly) agree that "one or more of the proposed designs would increase the probability of [them] reporting [their] symptoms", while 27.5% (strongly) disagree with the sentiment (see Figure 3.7). The responses amount to a mean of 3.37 on the Likert scale. However, out of all participants 60.0% claimed they would use the standard application at least weekly in any case, while only 18.0% stated they would never use it (see Table 3.1), meaning that, in theory, 60.0% of the people asked would not need any additional incentives to report their symptoms and which explains the fairly low agreement rate. Given the amount of people participating in ILI monitoring systems, this does not seem to be representative of the actual amount of people willing to report their symptoms regularly. For instance, in the Netherlands – the most active influenzanet region – only

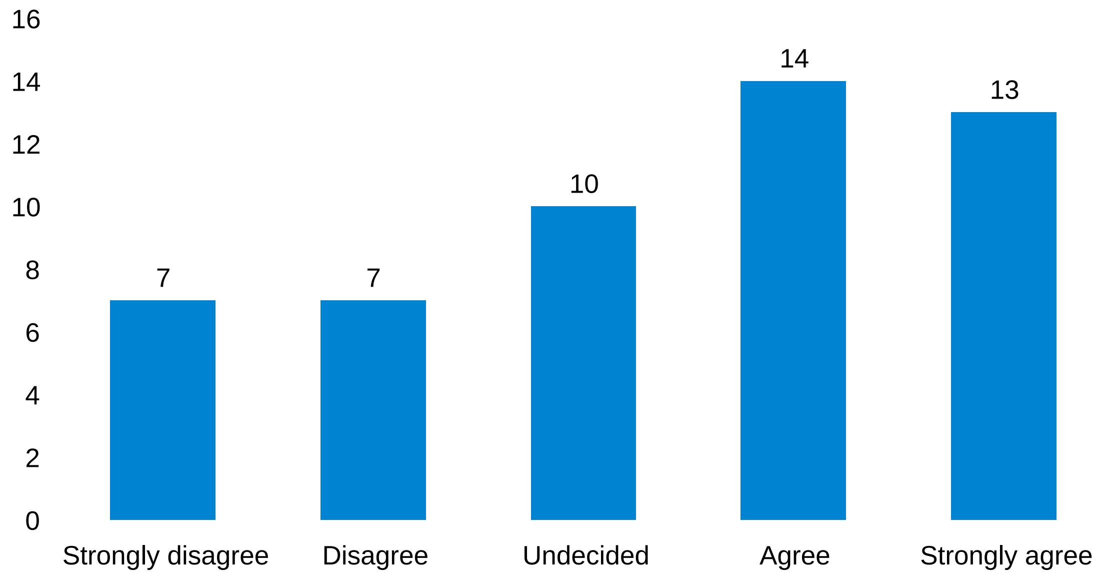



**Figure 3.7:** Amount of respondents agreeing to the statement that "one or more of the proposed designs would increase the probability of [them] reporting [their] symptoms". A five-point Likert scale was used as a scaling response approach.

| | Standard application | Gamification | Storytelling | Integration of SNSs | Diagnosis | Fitness correlations |
|---|---|---|---|---|---|---|
| Never | 9 | 15 | 21 | 26 | 12 | 14 |
| Monthly | 11 | 5 | 10 | 13 | 14 | 5 |
| Weekly | 27 | 22 | 15 | 11 | 23 | 18 |
| Daily | 3 | 9 | 5 | 1 | 2 | 13 |
| Several times a day | 0 | 0 | 0 | 0 | 0 | 1 |

**Table 3.1:** Amount of respondents stating with which frequency they would use a particular designed application.

12,000 people report on average, which amounts to about 0.071% of the population [Inf16]. This is most probably not solely attributed to a lack of knowledge.

Figure 3.8 shows which of the suggested designs would be an improvement according to the respondents. Out of all proposed designs, rewards in general were received best with 70.6% of respondents thinking that the integration of activtiy trackers would be beneficial and 54.9% agreeing that a diagnosis tool would improve participation.

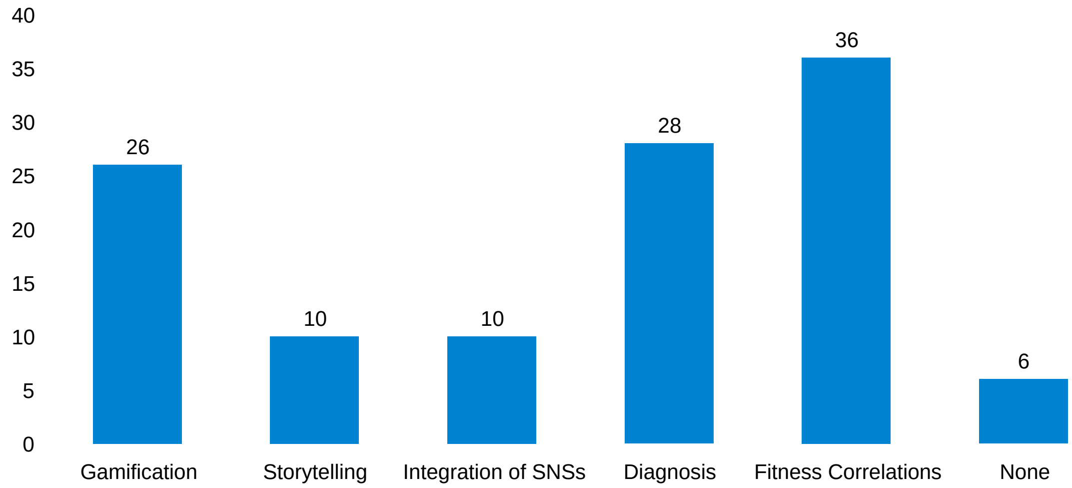



**Figure 3.8:** Number of respondents voting on the design proposals that would be a good improvement in their opinion.

According to the survey, the integration of activity trackers would also increase the frequency with which the application would be used. As seen in Figure 3.1, 25.5% of participants claim they would use it on a daily basis in contrast to 6.0% with the standard design. However, some concerns were raised for users who do not own or use an activity tracker or application, since 27.5% also stated they would never use the application (compared to the 18.0% never using the standard application). However, not owning an activity tracker would not change the basic functions of the application. Results also show that both, the diagnosis tool and the fitness design would be primarily used when the user feels sick (68.6% and 45.1%, respectively) and during flu season (39.2% and 31.4%). While not optimal, these designs should therefore be able to acquire the necessary data to detect disease spreading. Some survey participants also suggested that other rewards would be beneficial for participation, such as information about health care but also coupon codes, amazon coins and air miles. However, monetary rewards in general seem controversial, with 68.7% being either indifferent or disagreeing with the sentiment that they would be beneficial. They are also not the focus of this research.

The gamified approach reached the third highest acceptance rate with 51.0% of participants thinking that it would improve participation. Similarly to the fitness design, gamification would also improve the usage frequency with 17.6% of users reporting they would use it daily but simultaneously discourage many users from using the application at all since 29.4% reported they would never use it (see Table 3.1). It might also change the target demographic (users should usually consist of random citizens), since especially older people would be less interested in playing a game [GDC04]. Additionally, the majority of participants (54.9%) claimed they would use a gamified design mostly when they are bored or when they need to fill the time (e.g. waiting for the train, bus, etc.) (41.2%), which might not be desirable for an application that requires regular (weekly) participation.

The design with storytelling and the design with social network integration were both received as the least beneficial methods with only 19.6% of participants thinking they would improve participation. It could be reasoned that this is at least partly due to privacy concerns, as 59.2% of participants stated they have not used some mobile applications before because of such concerns. This goes especially for social networks where it was mentioned multiple times in a free text field for possible problems. A design that implements storytelling also seems to be mostly used when the user is bored, according to 47.1% of participants, which, again, is not exactly desirable for data collection on a regular basis.

Only 11.8% of respondents reported that none of the suggested designs would improve participation in their opinion.

Some general concerns were voiced, mostly with regard to privacy but also with regard to user experience as more functions might confuse users and make them turn away from an otherwise easy to use application.

### 3.3.2 Discussion and Conclusion

The online survey shows that some of the suggested designs may offer good incentives for voluntary data collection. Even though most respondents claimed they would use the standard application on a weekly basis anyway, they predominantly reported that an addition of the suggested incentives would nevertheless improve the application and therefore increase the usage rates.

Generally, designs that reward users with valuable information or statistics were received best. People seem to be interested in designs that provide tangible rewards that can be applied in real world circumstances. It can be reasoned that this is because the intrinsic purpose of the application changes. Instead of having an ILI reporting app with weekly surveys, users have a diagnosis tool or a personal health and activity tracking app. From that point of view, reporting symptoms becomes merely the necessary step to use the application's functionalities. According to the online survey, though, monetary rewards would also be received well which coincides with previous studies.

Gamification also looks promising but exhibits issues with usage frequency and acceptance rate in certain demographics. This can be explained by the additional effort and learning curve accompanied with acquiring the necessary skills to learn a game. Likewise, a gamified design makes the application look unnecessarily complex in view of the actual required functionalities for a data collection tool. However, for tasks that require human intelligence and that entail a certain skill that can be learned, gamification could prove to be very effective if the design exceeds the simple points/levels/badges reward system.

SNSs integration and storytelling proved to be the least well received designs. The former is not especially surprising, considering the associated privacy risks and issues. It can be argued that some demographics are more willing to expose themselves to such risks than others. It should also be mentioned that the risks involved highly depend on the data that is collected. In general, health data is seen as very private information that people are usually not willing to disclose to others.

The poor reception of storytelling can be explained by the lack of actual stories. While stories fascinate and often captivate people, they are usually consumed via novels, movies or games. Personalized messages and a journal might just not be enough to truly engage the user. This assessment can be justified by the majority of respondents reporting they would use the application when they are bored (47.1%). The design could be expanded with additional elements like pictures or soundtracks that further elicit positive emotions in order to bond with the user. This could prove a lot more effective, considering visuals make up about 80% of our perception [Ass12] and the vast majority of people consume stories via movies or television and create them by posting pictures on timelines. Another benefit of storytelling is that it can be easily implemented as an additional, complimentary element in the application design.

Important to note is that the target demographic may change depending on the applied design technique. For instance, especially older people raised concerns about a gamified version, which in their opinion over-complicates the application. Similarly, integration of SNSs targets a younger audience that has fewer issues with privacy concerns. One has to be careful with designing such an approach. However, depending on the nature of the desired data, this can also be used to ensure a certain target demographic as userbase.

As the most popular design approach, an application that includes activity data seems to be the best choice for a prototype to then further examine its effectiveness as an incentive for data collection. It should be noted, however, that while this design is specifically aimed at health oriented data collection, the online survey revealed that these kinds of rewards (i.e. access to more information, statistics etc.) would be well received in general.

A prototype with activity and health statistics was implemented in the following and afterwards evaluated.

# 4 Implementation

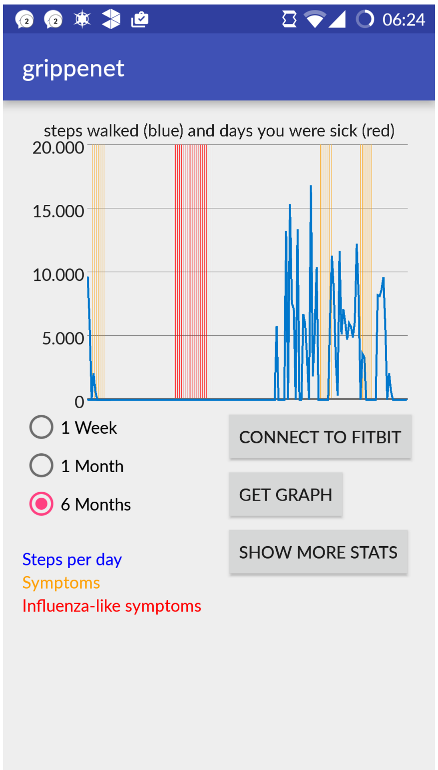
A prototype was designed in cooperation with the the German Research Center for Artificial Intelligence (DFKI) Kaiserslautern. A modular mobile application, temporarily called grippenet, to collect disease data via weekly surveys is currently being developed for Android at the DFKI as part of the CIMPLEX (Bringing CItizens, Models and Data together in Participatory, Interactive SociaL EXploratories) project. Thus, for the scope of this thesis a platform was provided with the necessary basic functions, including, on the user's end, a weekly symptoms survey, a heatmap that shows reports of flu-like illnesses and the possibility to easily include additional functionalities in future updates. In the course of this study, an additional module was implemented as an extension to the existing framework. As mentioned in the previous section, a module that processes activity data was chosen in order to correlate activity and fitness with the user's health data.

The existing framework is still under development but provides the user with a unique profile under which general user data and all survey results are stored on a private server at the DFKI Kaiserslautern. A survey module prompts the user each week to submit their symptoms. If the symptoms show a matching with typical flu symptoms, an ILI flag is set.

The following section explains the implemented module, from now called activity module, from the user's perspective.

## 4.1 The Activity Module

At this time the module requires the usage of a Fitbit activity tracker in conjunction with a Fitbit account in order to obtain personal fitness data. The user can track their activity with a given Fitbit tracker, synchronize it with their online account and retrieve the data to the grippenet application. The grippenet application itself stores the necessary health data, i.e. the weekly reported symptoms, on a server at the DFKI which can be accessed by the application.




(a) Main screen with buttons for navigation and the health data shown as bars overlaid with a line plot of the activity data.


(b) Bottom part of the secondary screen with additional statistics shown as line plots in addition to the computed mean scores.

**Figure 4.1:** Fitness module with main screen and secondary screen showing data for the last 6 months.

A Fitbit Alta Fitness Wristband [Fit16] was used to collect activity data over a period of roughly two months. The Alta Wristband tracks steps and distance walked, burned calories, sleep, diet with calorie intake, active minutes (sports) and non-dormant hours. All of these types of data can be retrieved and implemented in the module. However, for this prototype only steps, distance and calories burned were taken into account.

The user can access the statistic parts of the application via the menu located at the top of the application. From there, all other modules, e.g. the survey and map, are also available. Once the user is on the main statistic screen, the application will automatically connect to Fitbit and retrieve the necessary data. If this is the first time the user connects to Fitbit, a browser window is opened with Fitbit's web address where the user is asked for the login data and afterwards for the necessary permissions.

After the user connected to Fitbit, a diagram is shown on the main screen with bars for sickness reports and a line plot for the steps taken in the chosen timeframe (see Figure 4.1a). The steps are shown on the y-axis, while the x-axis shows the point in time. The bars are only shown if there was a report (y-axis) at the given point in time (x-axis) and are yellow if there was a sickness report that does not match influenza or red if it was an ILI. The timeframe can be changed with radio buttons (1 week, 1 month, 6 months). In case the Fitbit data did not get retrieved correctly, the user has a choice to try again by pressing a button and to then load the diagram data.

A second screen shows additional statistics (see Figure 4.1b). Again, the data is mainly visualized by plots. On the top of the screen, the bar diagram with the sickness reports is shown for the time frame chosen on the main screen. Beneath the diagram different line plots are shown for Fitbit's data and in the same timeframe in order to easily see correlations. From top to bottom, the visualized data are steps taken, distance traveled and calories burned. All of the plots are zoomable and scrollable so the user can adjust them to their liking. On the bottom of the screen, some general textual statistics are shown. These include mean scores like average steps walked per day, average distance traveled and sick days in the given timeframe. For these mean scores, days that had 0 activity were not taken into account, because a steps taken or distance traveled score of 0 suggests that the user did not wear their activity tracker that day which would then falsify the mean scores.

The application was implemented in Android using Android Studio 2.1. It was tested and used on a OnePlus One using Android 6.0 (Marshmallow). The following section provides an explanation of the implementation.

## 4.2 Implementation of the Module

Android was chosen because it is the most distributed operating system for mobile devices, making up about 82.8% of all mobile operating systems, followed by iOS (13.9%) and Windows (2.6%) [IDC15]. About 98.4% of all Android devices are using API 15 (Ice Cream Sandwich) or higher [Inc16a]. Since it is also sufficient for all our needs, it was chosen as the target API. Android was also chosen because the DFKI is working on an Android aplication that gathers illness data from users in order to help with the prediction of influenza or ILI spreading. The currently being developed grippenet application uses a modular structure that allows easy addition or modification of modules (functions) without altering the existing functions. It therefore serves as a framework for which a module with the intended functionality, as explained above, can be implemented. Figure 4.2 shows the fitness module that had to be implemented into the existing grippenet application. Mainly a GUI and the necessary computations to plot

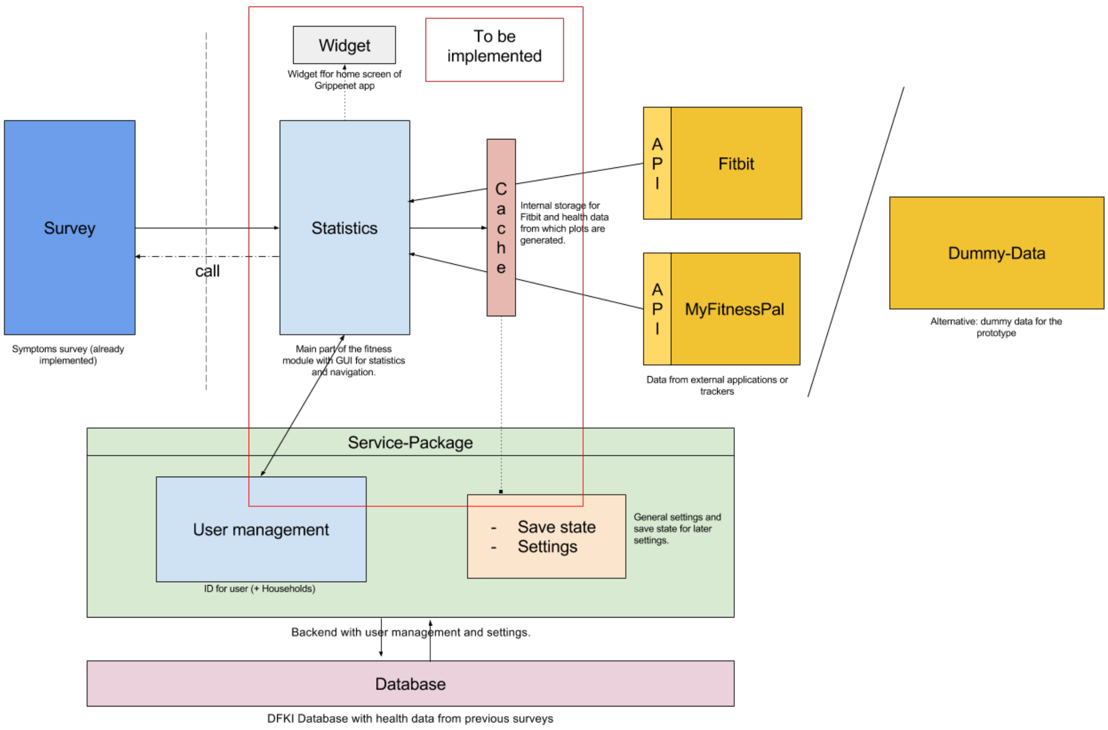



**Figure 4.2:** Structure and concept of the grippenet application with its data and communication layers. The marked section had to be implemented into the existing framework.

data had to be implemented. Additionally, a communication layer for the required data exchange with the desired fitness tracking service had to be created.

Fitbit was chosen as an activity tracking service since it is widely used and provides enough activity data to create insightful statistics. Fitbit also provides an API using OAuth2 which enables developers to easily obtain the user's gathered data as activity logs.

The module has two screens. The main screen shows one plot, buttons for navigation and a radio buttons to choose the desired time frame. With the buttons it is possible to access Fitbit's API. If this is the first time this is done, an additional browser window is opened where the user has to authenticate by providing the necessary login data. The user has then to allow the app to gain access to the necessary data in which case Fitbit sends an access token for future requests. Once done, the application can authenticate with said access token. The authentication itself is done with the OAuth2 protocol. The application connects to the Fitbit homepage and is then redirected to the authorization page. After that, Fitbit gives an authorization code to the application's callback URL.

The authorization code can then be exchanged for an access token and a refresh token (see Figure 4.3).

Once an access token is present, data can be retrieved anytime from Fitbit by sending the access token with a request. In the application, this is done automatically when the fitness module is accessed but can also be done manually by pressing a button. A timeframe and the desired type of activity has to be chosen for the request, whereupon Fitbit sends the desired data for the given timeframe as an activity log in a JSON format.

The activity log is filtered for the desired data

The filtered data is visualized as multiple plots, most of it on the secondary screen which shows the majority of the statistics (see Figure 4.1b) and can also be accessed from the main screen via pressing a button. The plots are calculated using GraphView for Android, an open source graph package that allows plots and diagrams to be generated from a given dataset.

Fitbit's activity data is visualized via line plots. These are clear and easy to understand and particularly sensible for data points like steps, distance, calories. Since no other data, like e.g. diet, was used in this prototype, line plots are perfectly suited for the purpose of this application and can easily be compared to each other.

For the sickness data, however, bar diagrams were used. This is because the vsualized sickness data is binary (either there was a report or there was none) or, at most, tertiary, if we differ non-ILI reports from ILI reports. A bar is shown when there was a report that week and no bar if there was no report, respectively. The difference between non-ILI and ILI reports was realized with different colors. A yellow bar shows a report that does not match influenza like symptoms and a red bar shows if there is an ILI flag present on that week.

The sickness data itself is retrieved from the DFKI servers. It is previously stored with the user's profile and, similar to the Fitbit data, can be obtained in a JSON format after authenticating to the server. However, since the application is still under development at the DFKI Kaiserslautern and the timeframe for this study is limited, no surveys could be taken to accumulate the neccessary data. Instead, dummy data had to be used for the purpose of this study. The dummy itself stems from the author's personal and potential health condition for a timeframe of about three months.

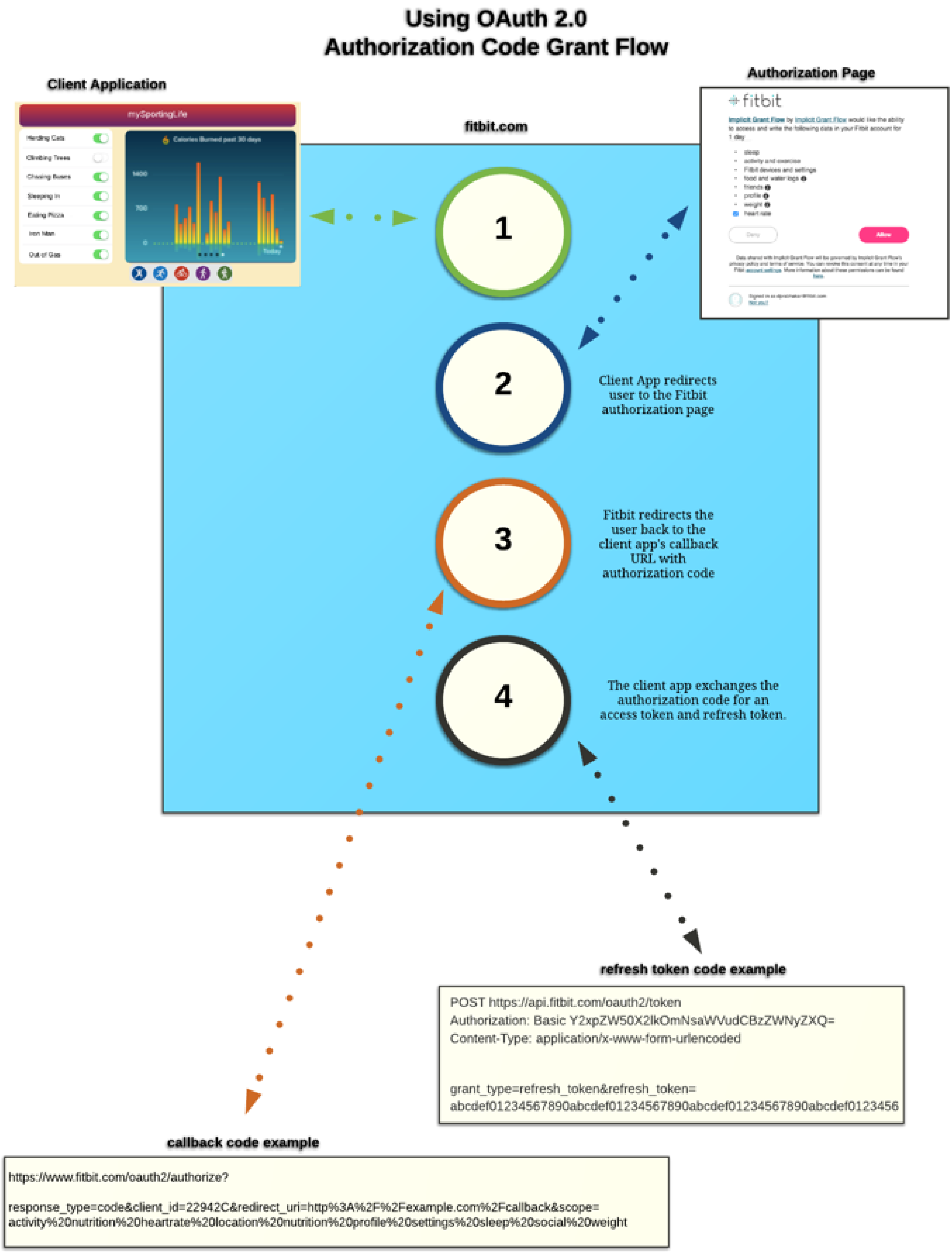



**Figure 4.3:** The OAuth2 authorization protocol with Fitbit. Source: fitbit.com

# 5 Evaluation

Given the prototype from the previous chapter, a user study was conducted using ContraVision [Man+10] to evaluate the idea of combining fitness and health. A user study with ContraVision introduces a new technology via short videos. Typically, there are two videos about the new technology or software: one video showing only positive aspects of the technology and one negative version of the same video that brings all the disadvantages and concerns to light. Each video consists of multiple scenes, whereby every scene starts with the same premise but develops into different directions, depending on the version. The positive video also usually shows mostly positive emotions and feelings while the negative version focuses on showing negative connotations like doubts or concerns with the technology.

The videos are then shown to a small group of people, usually in a domestic environment in order to avoid sterile and unauthentic environments. Afterwards, people have to answer a few questions in individual interviews. The interviewees are then encouraged to discuss their perceptions and thoughts in a group in order to get to a mutual conclusion.

ContraVision is a good and simple approach to test new technology (software or hardware) with respect to its influence on current society. It is also very easy to estimate how well said technology will be received and accepted by people. Since this study focuses especially on acceptence rate and willingness to participate in data collection, ContraVision appears to be a good tool for evaluation, as it provides very good insight into these aspects, particularly because it does not rely on a sterile and controlled environment.

## 5.1 User Study

Two videos were filmed using the storyboards shown in table 5.1: one where the positive aspects of the application are shown and one where the negative aspects are brought to light. In scene one, the main character Michael is introduced to the application by a good friend, in particular to the ability to connect his new fitness tracker with the app in order to see connections between his health and his fitness. In the positive version,

| **Positive version** | **Negative version** |
|---|---|
| *Scene 1: at a friend's house* | |
| Michael shows his new fitness tracker and expresses curiosity about the correlation between his fitness and his health. His friend suggests to use Grippenet and explains what it does. Michael asks questions and seems *intrigued* by the idea. He *optimistically* agrees to test the application. | Michael shows his new fitness tracker and expresses curiosity about the correlation between his fitness and his health. His friend suggests to use Grippenet and explains what it does. Michael asks questions and seems *sceptical* about the idea. He *reluctantly* agrees to test the application. |
| *Scene 2: Running* | |
| Michael runs with his fitness tracker and records data. He looks at the statistics for the last month and seems satisfied with the results. | Michael runs with his fitness tracker and records data. He looks at the statistics for the last month and seems unsatisfied with the results. |
| *Scene 3: at Julian's home* | |
| Michael *enthusiastically* tells his girlfriend about Grippenet. He shows her the basic functions of the application and she seems *excited* about it and *looks forward* to future results. | Michael tells his girlfriend about Grippenet. He shows her the basic functions of the application and she seems *doubtful* about it and its future results. She also expresses *concerns* about privacy with which Michael agrees. |
| *Scene 4: on the street with Julian* | |
| Michael shows his friend the results (statistics) of the last few months and tells him that he felt comepelled to be active. He wants to keep using Grippenet to see his future performance and health. | Michael shows his friend the results (statistics) of the last few months and tells him that he didn't feel like he saw any interesting changes. He will not keep using Grippenet as the information doesn't really help him. |

**Table 5.1:** Storyboards for the positive and negative video, using the same scenes with different outcomes in order to evaluate the prototype with ContraVision.

Michael seems interested and willing to try the application, while in the negative version Michael shows doubts about its usefulness and only accepts to try it out after some severe hesitation (see Figure 5.1). In scene two, Michael is shown running with his activity tracker and afterwards checking the statistics on his smartphone. In the positive version he seems satisfied and positively surprised. In the negative video he seems confused and disappointed. In the third scene, Michael tells his sick friend about this app he has been using for a while. In the positive video, he recommends the app thinking it would be helpful and useful for Julian's health. In the negative video, Michael still recommends the app but is soewhat hesitant to boast about it and seems doubtful about its usefulness. So does Julian. Julian also questions the privacy to which Michael doesn't have satisfying reply. In the last scene, Michael and Julian meet again on the road and talk about their experience with the app during the last few months, In the positive version, Michael claims to have seen some pretty interesting correlations between his fitness and his health while in the negative version he mentions that he forgot to participate in the survey quite a lot of times which is why he lost his motivation to participate at all. In both versions Julian agrees with the respective sentiment.

Additionally, in-character interviews were conducted, where Michael and Julian answer questions about their perceptions of each other and about their experience with the application. The answers vary, depending on whether the interviews followed the positive or the negative version of the video. The interviews were recorded with a microphone as sound files only.

The videos were shown to 10 people overall, divided into 4 different groups. Two of those groups got to see the positive version and the other two got to see the negative one. After watching the video, the participants listened to the corresponding in-character interviews of both Michael and Julian. The videos were about three to four minutes in length and so were the interviews for each character, averaging to about 10 minutes of watching and listening. After the voice files, the participants were interviewed individually by asking questions about the video/interviews and the introduced application. The interviews were conducted in a seperate room, so that their replies would not affect the opinion of the rest of the group. For each participant the interview took about five minutes on average. After the interviews, the participants were asked to sit together again in order to discuss their individual opinions in a group session in order to exchange their thoughts and strengthen their opinion. Group discussions took between five to ten minutes, depending on group size and consensus within the group.

**(a)** Section from the positive video. Michael (to the left) is excited after hearing about the application and actively participates in the conversation.

**(b)** Section from the negative video. Michael (to the left) is stressed and slightly irritated after hearing his friend's explanation.

**Figure 5.1:** Screenshots from the first scene of both videos. The scene is showing Michael and his friend who is currently explaining the grippenet application to Michael

## 5.2 Results

The interviews gave a wide range of different aspects concerning the app in general. By comparing not only the positive interviews witch each other and, of course, the negative interviews respectivly, but also comparing the positive with the negative interviews, one can find similiarites on both sides. To structure these findings, first the the positive interviews will be looked at and especially arguments that were voiced more than once will be considered. The same will then be done for the negative interviews. After that, the positive and the negative interviews will be compared and examined for reoccuring arguments. Finally, conclutions will be drawn out of these findings to see how the interviewees saw the app in general.

### 5.2.1 Evaluation of the positive interviews

The positive interviews show one great general consensus: the application, per se, is regarded as helpful and the usage is not only good for yourself but also for others. Multiple interviewees praised the functionalities of the application:

> *"... the app seems to be competent and does what it is supposed to do."*

> *"... the app gives an interesting possibility of comparing your fitness with your health."*

In accordance with these statements, all interviewees reported that they would use the application themselves.

A frequent topic that was brought up, is the security of the transmitted information which is not especially surprising, considering the nature of the data that is collected. However, opinions differ in this aspect. Only one of the interviewees voiced actual concern about personal health data.

> *"One should think about the security of the data."*

The other interviewees did not seem to have any issues with the the submitted data. One of them simply said that

> *"I see no problem in disclosing this kind of information."*

while another responded that

> *"The transmitted data is not really that significant, therefore I see no problem in using the application."*

So even though there are some concerns, it is not predominantly considered as a privacy issue, because the amount and the type the data actually being transmitted through the application does not reveal any major personal aspects about oneself.

A few other issues were observed, however. For one, an activity tracker is needed to get information about one's personal fitness, i.e. in order to see the correlations between one's health and one's activity. This concern is voiced by one of the interviewees, who says that

> *"I don't own a fitness tracker and I don't really walk or jog that much. The overall usage therefore is difficult for me. If I had an activity tracker I would use it in my everyday life."*

The quote also hints at the second issue. In order to get any meaningful results, you have to do at least some activity. If you do not participate in any kind of activity which is out of the ordinary (e.g. sports) your activity data will most probably be within a very narrow margin which makes noticing useful correlations very difficult.

This sentiment can also be observed in one of the other interviews, where the respondent replied that:

> *"The concept to see the relation between your activities and your health is interesting. But I don't know if I would use it myself, because I am not really into any sport."*

Another important factor to consider was revealed in the group discussion. Here, interviewees addressed the problem of long-term usage of the application. They raised the question whether motivation to use the application lasts long enough to see some meaningful correlations in the generated statistics.

> *"I would test the app but perhaps I wouldn't always use it or wouldn't use it for a long period of time."*

Finally, there is the matter of recommending the app to other people. Here the interviewees have different opinions. One says that

> *"I would basically recommend it to others."*

while another said that

> *"I probably would not recommend it. I am more of a quiet type and would only use the app for myself. I would use it and only then recommend it if I saw any specific reason to do so, like really helping others. However, I like having someone to motivate me to do more sport."*

Lastly, one interviewee said that

> *"It depends on knowing someone who would be interested in an app like that but in principle, I would recommend it."*

Interviewees basically agree that they would only tell others about it if they have seen a positive effect after using it for some time themselves. They would also get a better grasp of the app itself which makes it easier to discern who might be interested in using it and who might not be.

### 5.2.2 Evaluation of the negative interviews

Following the same structure as for the positive interviews, an overview of the general reception of the application will be given. In the case of the negative interviews the interviewees do not seem to come to a uniform opinion.

Generally, interviewees seemed to be reluctant about the application and especially about using it themselves:

> *"If enough people participate in using the app, then I like the general idea. But if there are not enough participants, the app doesn't make any sense. I think that in the long run the app isn't going to be used by many."*

A similar opinion was voiced about Michael:

> *"I would have tested the app in more detail, [...] and would have only used it regularly if I deemed it useful."*

One interviewee, in particular, seemed to be opposed to the application in general. According to them, the application does not add any useful information:

> *"I don't know the app well enough but I probably wouldn't use it because I can't really see a personal benefit. I would prefer it that the app gives me inforamtion and tips, when I am already sick, to get better."*

Another interviewee offered a completely different perspective. On one hand they said that

> *"I would test the app for myself. Then, if I think it's useful, I would continue using it and even recommend it to others."*

On the other hand they voiced major concerns:

> *"Honestly, the heatmap rather scares me than being useful to me. Espacially when you see where a lot of reports of sick people are. From this information, people would be scared to go to certain areas or cities, which I think is problematic."*

An important observation to make about these interviews is that even though interviewees seemed somewhat reluctant about using the application, it was nevertheless received as an interesting concept that most people would want to use or at least try.

Some concerns were voiced regarding actual usefulness in everyday life, however. An interesting observation to make is that most of these concerns were raised as questions. The two main ones were: "Are the statistics really meaningful?" and "Will enough people participate to atually give any meaningful results?"

Surprisingly, the interviewees did not seem as concerned about their privacy as Michael's friend. Even though health data was perceived as very personal information, the interviewees seemed to be willing to give it away for a public good.

> *"I don't have any problems to give away this personal information for a charitable cause. I simply don't want my data to be used for anything else."*

Some even reported that they do not care at all about their personal information as long as it is collected anonymously.

> *"I would use it. The data about me is anonymous, that's why I don't really see a problem. I don't care about talking about my personal health in public."*

One interviewee voiced genuine concern about their personal information:

> *"Yes, [...] I would like to know where my data is going, espacially who actually can see them and what happens with it."*

So even though interviewees were more reluctant when asked whether they would use it personally, for most respondents, privacy issues did not seem to be one of the reasons for it. One interviewee even observed that the whole point of the application is to collect as much data as possible in order to be meaningful.

The topic of doing any sport and or even having an activity tracker did not get mentioned in any of the negative interviews.

Instead, concerns about long-term motivation were raised mutliple times during the individual as well as group interviews. For instance, concerning Michael's handling of the application one interviewee said:

> *"I probably would have used the app more regualarly, but I fear that my particpation would drop over a longer duration."*

The same interviewee later continued by mentioning that:

> *"I think, in the long run, only a few will use the app."*

Another very similiar statement was made claiming that

> *"I would really only test it in the beginning. But over time I wouldn't really use it regularly anymore."*

Another interviewee who voiced the same concerns, also suggested a solution that would work for them.

> *"I would set up a reminder in my phone. By that I would have forced myself to use my fitness tracker as well as the app."*

It is clear that the long term motivation is considered a problem or at least called into question. Users might get frustrated by not getting immediate results and give up before meaningful patterns and correlations can be discerned.

The interviewees also seemed reluctant when asked how they would feel when they tell others about the application:

> *"I would test the app for a longer time for myself. Only if I think it is useful, then I would tell others about it."*

Another interviewee revoiced their privacy concerns which would make them reluctant to recommend it to others:

> *"I would test it for myself at first and only if I saw any results I would recommend it. But I am still unsure about the data security and the deterrent effect of seeing where a lot of cases of influenza are reported."*

According to the interviewees' replies, poeple would try the application but only tell others about it when they made sure that it is save to use as well as useful in general.

In summary, even though a negative video about the application was shown to the interviewees, they still reported that they would at least try the application. They seemed a lot less excited about it, however, since they did voice some concerns. Some interviewees raised questions about the data usage, i.e. considered their privacy to be of concern. The biggest doubts, howver, were mentioned about the lack of motivation for long-term usage. Neither the purpose of the application and the fitness statistics, nor the functionality was seen as questionable, per se. What made people reluctant was the time investment that would be necessary before meaningful correlations between one's health and one's activities can be shown.

### 5.2.3 Comparison of the positive and negative interviews

Both types of interviews show some similarities as well as differences that are shown and discussed in the following. This helps with getting a complete and somewhat unbiased picture of the application, in particular in regard to user's motivation to participate in the weekly surveys.

Both groups reported that they were willing to use the application. Interviewees that saw the negative version of the video showed more reluctance, though, and generally less excited. Nevertheless, this shows that people perceive the idea and the concept of the application as mostly positive and worth using.

The observation that people who saw the negative version seemed to be less excited about using the application themselves, could be explained by their concerns about long-term usage. More than once, doubts about long-term motivation were voiced, since the generated statistics only become menaingful after a long period of time. This is in line with Michael's experience in the (negative) video. He loses his motivation pretty quickly and sometimes forgets to use his tracker as well as the application itself, which is why the statistics do not give him any meaningful insight, even after a considerable amount of time. Interviewees who saw that version appeared to reciporate this sentiment.

Interestingly, especially interviewees who saw the positive version reported that Michael appeared to be too excited about the application and was even perceived as naive regarding this. This can be explained with Michael's acceptance rate for an application that uses personal information and his willingness to recommend it to others almost immediately. in the negative versions, he is asking a lot more questions about the application and seems to be hesitant to recommend it to his friend. Only in that version does he raise the issue about his privacy.

This tells us that people are generally not as willing as Michael to accept a new application which collects personal data about the user and that they think about the implications for their privacy, even if Michael does not.

Another important topic is the concern about ones privacy, which came up multiple times in the positive as well as the negative interviews, even though Michael only talks about the issue in the negative version. Surprisingly, both groups voiced mixed opinions about this topic with some interviewees being more concerned than others and some interviewees not being concerned at all. This suggests that people's privacy concerns are not dependent on the presentation of the application but rather on their personal opinion and maybe experience on the matter.

Regarding people's feelings about talking to others about the applications, there are some similarities but also differences in both types of interviews. Both do not have any problem with talking to others about it, per se, and both are also willing to recommend it to others eventually. However, interviewees that saw the negative version were on average more reluctant to recommend the application. Most reported that they would want to test the application themselves and learn about its usefulness, before they recommend it to others. Interviewees who saw the positive version were slightly more willing to share information about the application with others, even if they do not know its effectiveness yet. This sentiment is completely in line with Michael's behavior in the videos but it could also be attributed to personal preferences as the difference is minor and the sample size of interviewees relatively small.

Neither interviewees of the positive nor interviewees of the negative version were completely unwilling to tell others about the application or to recommend it. This, again, shows us that the general concept was appreciated – albeit sometimes with reservations.

### 5.2.4 Summary of the findings

Considering all interviews, the application was generally received well. The interviewees saw the potential of the application, its purpose and agreed that its inclusion of activity tracking could be meaningful. They also realized that a regular investment has to be made over a long period of time in order to get some insightful information. Therefore, some interviewees raised concerns about the long-term motivation. It stands to reason whether an improvement could be made by giving the users an immediate reward – even if it is only virtual such as points or badges.

Privacy concerns were brought up multiple times but usually not perceived as severe enough so that people would not use the application. This is mostly due to two reasons: For one thing, the gathered data is anonymous and not associated with any personal contact info that identifies the user. For another thing, people seem to be more willing to submit health specific data if it serves a public good.

People also voiced concerns about the requirement of an activity tracker as well as about them not doing any physical activity which might compromise the application's usefulness. Regarding the former, however, it should be said that Fitbit was only used for the prototype. Other approaches can be made with, for instance, Google Fit [Inc16c] (which is completely free for Android) or even with manual input. Regarding the latter, not being active in any way might indeed prevent the user from finding any meaningful correlations between their (consistently low) activity and their health. A point should be made, however, that the user might also be motivated to be more active in that case, because only then the application provides informative statistics. Additionally, it is possible to include one's diet or heartrate (if measured) as activity data which is something that can be used for every user, no matter what the physical activity amounts to. Of course, the applcation might be more insightful for people who are physically exercising.

In conclusion, the application is generally received well and people seem willing to use it themselves. However, in the course of this study, three major problems were discovered:

1. privacy issues with submitted health data
2. additional requirements that are necessary for insightful statistics
3. lack of motivation for long-term usage

The first problem was regarded as acceptable when the the data is anonymous and used for public good. The second problem can be circumvented by implementing other applications or activity data which every user has and can use without additional tools. The third problem is more difficult to solve. This raises the question whether it would be better and also feasible to include additional rewards that become available immediately after the user submitted their data.

# 6 Discussion

Section 3.3.2 showed that some mobile application designs were received rather well and would improve participation of potential users. The two reward-based designs were seen as the best incentives, while integration of social networks and storytelling received only moderate support in the online survey. The findings for each particular design are discussed in the following.

Gamification, while not the most popular design approach, can be a good design choice for general purposes, i.e. where tangible rewards are difficult or even impossible to implement. Gamification can be especially effective where human intelligence is required and if there is a skill to be learned, as reported by [Coo10].

Gamification is also one of the few design approaches that can provide true intrinsic motivation, i.e. pure enjoyment or curiosity which typically amount to a better long-term incentive than extrinsic motivational factors. However, the online survey showed that gamification might over-complicate the application which is why a lot of people might shy away from putting in the effort necessary to learn how to use it.

Storytelling proved to be not motivational enough to be a suitable design choice. Most people reported they would only use the application when they are bored.This suggests that the design itself is too boring to provide a true incentive. In the suggested design, however, only textual story elements were suggested and shown. It is worth pointing out that the inclusion of more media such as pictures, videos or sound could vastly improve a storytelling approach, especially because visual information makes up about 80% of our perception [Ass12].

Additionally, storytelling is one of the easiest designs to implement and also very simple to adjust to an existing design accordingly. It might therefore prove to be a good complement to other design techniques as long as it does not over-complicate the usage of the application – which was regarded negatively about the gamification approach.

Similarly, integration of social networks proved to be rather impractical. One possible and probable reason for this is the concern for one's privacy. In the online survey, 59.2% of respondents reported that they refrained from installing at least one mobile application before because of privacy concerns. Free text questions also showed that privacy is an important issue for people. In the prototype user study, interviewees revealed that privacy is indeed of concern to them but that they are willing to submit the data as long as it is anonymous. Sharing one's health on SNSs defeats that purpose, however, which is why considerably less people are willing to do it with that design approach.

It should be noted, however, that integration of SNSs could still be useful for marketing purporses. By sharing one's usage of the application or one's health, other people may become interested in the application and learn about its purpose which might attract new users and therefore considerably increase the quantity of gathered data. Flu Near You allows users to share their health on SNSs which is an indication of its usefulness as a marketing tool. However, how to market and how to spread knowledge about the application (or the data collection in general) was not the focus of this study. As an incentive to submit data, integration of SNSs is of little use,

Providing rewards seems to be the best choice to incentivize people to submit their data after all. Two designs were suggested: A diagnosis tool and a fitness activity tool. Both were received rather well with the latter drawing slightly more interest. However, the suggested approaches are both very application-specific which makes evaluating them on a grander scale difficult. It could also be reasoned that those approaches change the actual purpose of the application from a user's perspective. It is no longer a tool to collect data but a tool that processes and shows data generated from user-specific inputs. Changing people's perception of the application's purpose might be a good way to incentivize users as long as it is made clear that data is collected, what that data entails and what it is used for.

Even though the design was very application-specific, it can be concluded that people want some tangible reward or effect that they can actually use in real-world circumstances. That sentiment is stressed by both, the results from the online survey as well as the evaluation of the fitness statistics prototype. One interviewee of the prototype study even suggested to provide more tangible rewards that are available immediately.

It is not particularly surprising that tangible rewards are also received well, since previous studies reported that monetary rewards provide good incentives. It is interesting to see, however, that people seem to be more willing to provide their health data when there is a personal gain involved and when it is done anonymously and for public good. The latter was especially brought to light during the interviews but was also mentioned in the online survey with 56.9% of respondents reporting that contribution to public good is motivating for them.

A very important issue with the diagnosis tool is its intended purpose. It is intended as a tool that provides potentially helpful information by warning the user of possible infections. However, it could be abused as an actual diagnosis and be used as a substitute for a visit at one's physician which involves potential health risks, especially if the diagnosis is wrong. Distributing such a tool could also prove to be very difficult from a legal perspective, depending on the country or state where it is supposed to be released.

The integration of an activity tracker on the other hand, is simple to do and in retrospective an effective method to get people to care about their health and fitness. The online survey showed a large interest for correlations between activity and health which is why it was chosen to be implemented as a prototype.

The prototype merely used the most rudmimentary activity data such as steps and distance walked as well as calories burned. Measurements were taken via a Fitbit Alta wristband. It is possible to expand on this concept fairly easily with additional data such as one's particular diet, sleep schedule or heart rate.

The evaluation of said prototype was done via ContraVision, an evaluation approach that uses two similar videos, one showing positive and one showing negative aspects of the given device. The interviews showed that the concept was generally received well and that people would be more willing to use the application and submit their data than without the fitness statistics functionality.

Nevertheless, the evaluation process revealed three main issues with the application. For one, people care about their personal health data and are suspicious about their privacy. It was shown, however, that people are generally more willing to submit their data if it is anonymous, for public good and if the purpose is made clear. This is ensured with the introduced application as the only data that is collected is health-related and its purpose serves a charitable cause: predicting ILI spreading.

Secondly, some complained that many people do not own or a use an activity tracker or are not very active to begin with. It has to be noted that this is an application-specific issue that does not apply to reward based designs in general. It can still be improved, however, by using software instead of trackers such as Google Fit and using data such as diet, sleep schedule or heartrate where people do not have to live an active lifestyle in order to get useful data.

The third problem is the questionable long-term motivation that the suggested design provides. People only get meaningful statistics after a very long period of time. The question is whether people who are initially interested in the concept stay motivated long enough to use the application on a regular basis until they can get meaningful statistics. This might be a problem in many reward-based design approaches that do not use monetary compensation but rather grant access to new information, data or functionalities. This is also a limitation of this work as long-term motivation could not be evaluated in the given timeframe. A possible solution for future reward-based design could be to provide multiple different rewards – some more immediate than others.

Another important limitation of this work is the data quality. The purpose of this study was to motivate as many people as possible to submit their data via (mobile) data collection tools, i.e. to increase the quantitiy of datasets. Qualitiy of data has not been measured. However, at least for the suggested reward based designs, it can be expected that data quality is fairly high, since people would not get meaningful results, neither of the diagnosis tool nor of the fitness tool, if they do not submit thoughtful and completely honestly. For all other suggested methods, quality of data could potentially suffer and should therefore be investigated in future work.

This work was intended to provide a guideline of what design methods could work for (mobile) data collection tools in general. However, the suggested designs were very specific to flu symptom surveys. Even though this might prove to be a limitation of this study for future work and designs, it is still possible to draw general conclusions about what motivates people to provide their data. It is also clear that every data collection tool should be adjusted according to its intent.

Nevertheless, it is possible to draw general conclucsion about what motivates people to submit personal data. We have seen circumstancial factors that way heavy into people's willingness such as the intended purpose (marketing, public good, etc.). We have also seen that privacy plays an important role. People are more willing to submit their data if it is anonymous. Methods for designing an application to collect data can not change these factors. However, we have seen that some designs are more effective than others. According to the online survey, people are generally interested in any kind of tangible reward – not only monetary rewards. In order to create such incentives, application-specific rewards that match the application's intended purpose have to be provided such as access to more (relevant) information or functionalities. This also creates the possibility of changing the user's perception of the application: submitting one's data is only the requirement to get the necessary functionalities and is therefore perceived as less bothersome. This might be a good way to incentivize people but should be done with care, i.e. users have to be informed that their data is collected and for what cause.

Gamification has a similar effect of changing the application as a whole. It should also be considered as a design choice as it was received fairly well in the online survey and its potential is highly dependent on the implemented game. It is especially suitable for tasks where a skill that can be learned is required. It is, however, the most implementation-heavy approach which might make it inappropriate by default.

# 7 Conclusion and Future Work

Methods to motivate volunteers to participate in data collection were explored in this work. Four different techniques to incentivize users of a mobile application introduced and five prototypes for possible designs were investigated via an online survey.

The survey revealed that people are generally interested in rewards, especially if they serve a tangible purpose. Those rewards do not have to be of monetary nature but could instead include access to additional content, information or functionalities. While gamification proved to have potential as well, integration of SNSs and storytelling were not received well enough to be truly considered for future designs.

As reward based designs were received the best, a prototype of a fitness module was implemented into a framework for an ILI symptom survey application that is currently being developed at the German Research Center for Artificial Intelligence. The module uses Fitbit for activity data and tries to show correlations between the user's activity and health.

The prototype was then evaluated with contravision, a userstudy that incorporates two similar videos showing the negative and positive aspects of the application, respectively. The subsequent interviews showed that people liked the suggested prototype and considered it an improvement to the standard application that has no additional functionalities to the survey itself. All of the respondents were willing to at least try the application. However, the interviews also revealed that people care about their privacy and are concerned about long term motivation.

In summary, this work showed that data collection tools can definitely be improved through various design methods. Especially rewards proved to be appreciated the most so that more people feel motivated to submit their data.

Since the implemented prototype only yields valuable statistics after a longer period of time, long-term motivation could not be observed in the course of this work. It also constitutes a limitation of this work and should therfore be further investigated in general.

The quality of the collected data is another important aspect to consider for future work. The purpose of this study was to increase quantity of data – quality was not regarded.

This is especially important for designs taht are not dependent on the submition of correct data (gamification, storytelling, integration of SNSs, rewards that do not use the submitted data, etc.).

Another potential issue to consider is the work that has to go into designing and implementing the suggested methods. The cost of a gamification design, for instance, could prove to be considerably higher than one that integrates SNSs.

This work should provide an overview of possible design methods that can increase contribution to participatory data collection. It may provide a starting point to more thorough investigation regarding participatory data collection, which is an area that plays an important part in all research.

# Bibliography


[AMM11] E. Aramaki, S. Maskawa, M. Morita. “Twitter catches the flu: detecting influenza epidemics using Twitter.” In: *Proceedings of the conference on empirical methods in natural language processing*. Association for Computational Linguistics. 2011, pp. 1568–1576 (cit. on p. 21).

[AN95] M. G. Angur, R. Nataraajan. “Do source of mailing and monetary incentives matter in international industrial mail surveys?” In: *Industrial Marketing Management* 24.5 (1995), pp. 351–357 (cit. on p. 27).

[Ach+11] H. Achrekar, A. Gandhe, R. Lazarus, S.-H. Yu, B. Liu. “Predicting flu trends using twitter data.” In: *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*. IEEE. 2011, pp. 702–707 (cit. on p. 21).

[Ana16] S. Analytics. *SurveySwipe*. 2016. URL: http://www.surveyswipe.com/ (visited on 09/12/2016) (cit. on p. 20).

[Ass12] N. N.-O. R. Association. *Vision Is Our Dominant Sense*. 2012. URL: http://www.brainline.org/content/2008/11/ (visited on 09/12/2016) (cit. on pp. 37, 57).

[BKG11] M. Buhrmester, T. Kwang, S. D. Gosling. “Amazon’s Mechanical Turk a new source of inexpensive, yet high-quality, data?” In: *Perspectives on psychological science* 6.1 (2011), pp. 3–5 (cit. on p. 19).

[BMZ11] J. Bollen, H. Mao, X. Zeng. “Twitter mood predicts the stock market.” In: *Journal of Computational Science* 2.1 (2011), pp. 1–8 (cit. on p. 26).

[BSM12] J. L. Boyles, A. Smith, M. Madden. “Privacy and data management on mobile devices.” In: *Pew Internet & American Life Project* 4 (2012) (cit. on p. 20).

[Bre12] B. Brett. “The psychology of sharing. Why do people share online.” In: *New York Times, New York.* 22 (2012) (cit. on p. 26).

[Bur+06] J. A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, M. B. Srivastava. “Participatory sensing.” In: *Center for Embedded Network Sensing* (2006) (cit. on p. 20).

[Cha+12] G. Chatzimilioudis, A. Konstantinidis, C. Laoudias, D. Zeinalipour-Yazti. "Crowdsourcing with smartphones." In: *IEEE Internet Computing* 16.5 (2012), pp. 36–44 (cit. on p. 19).

[Chu+12a] R. Chunara, S. Aman, M. Smolinski, J. S. Brownstein. "Flu Near You: An Online Self-reported Influenza Surveillance System in the USA." In: *Online Journal of Public Health Informatics* 5.1 (2012) (cit. on p. 21).

[Chu+12b] R. Chunara, V. Chhaya, S. Bane, S. R. Mekaru, E. H. Chan, C. C. Freifeld, J. S. Brownstein. "Online reporting for malaria surveillance using micro-monetary incentives, in urban India 2010-2011." In: *Malaria journal* 11.1 (2012), p. 1 (cit. on p. 21).

[Coo10] D. Cook. *Ribbon Hero turns learning Office into a game*. 2010. URL: http://www.lostgarden.com/2010/01/ (visited on 09/12/2016) (cit. on pp. 25, 57).

[DM15] K. Dergousoff, R. L. Mandryk. "Mobile Gamification for Crowdsourcing Data Collection: Leveraging the Freemium Model." In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM. 2015, pp. 1065–1074 (cit. on p. 24).

[Det+11] S. Deterding, D. Dixon, R. Khaled, L. Nacke. "From game design elements to gamefulness: defining gamification." In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. ACM. 2011, pp. 9–15 (cit. on p. 23).

[Dug+13] A. F. Dugas, M. Jalalpour, Y. Gel, S. Levin, F. Torcaso, T. Igusa, R. E. Rothman. "Influenza forecasting with Google flu trends." In: *PloS one* 8.2 (2013), e56176 (cit. on p. 21).

[EHW03] J. U. Espino, W. R. Hogan, M. M. Wagner. "Telephone triage: a timely data source for surveillance of influenza-like diseases." In: *AMIA*. 2003 (cit. on p. 21).

[ESL06] N. Ellison, C. Steinfield, C. Lampe. "Spatially bounded online social networks and social capital." In: *International Communication Association* 36.1-37 (2006) (cit. on p. 26).

[Fer+05] N.M. Ferguson, D.A. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, D. S. Burke. "Strategies for containing an emerging influenza pandemic in Southeast Asia." In: *Nature* 437.7056 (2005), pp. 209–214 (cit. on p. 14).

[Fit16] Fitbit. *Fitbit Alta wristband*. 2016. URL: https://www.fitbit.com/alta (visited on 09/12/2016) (cit. on p. 40).

[Fla+11] D. R. Flatla, C. Gutwin, L. E. Nacke, S. Bateman, R. L. Mandryk. "Calibration games: making calibration tasks enjoyable by adding motivating game elements." In: *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM. 2011, pp. 403–412 (cit. on p. 24).

[GDC04] M. D. Griffiths, M. N. Davies, D. Chappell. "Demographic factors and playing variables in online computer gaming." In: *CyberPsychology & Behavior* 7.4 (2004), pp. 479–487 (cit. on p. 36).

[GIS16] GIS. *GIS Cloud Mobile Data Collection*. 2016. URL: http://www.giscloud.com/apps/mobile-data-collection (visited on 09/12/2016) (cit. on p. 20).

[Gao+11] H. Gao, G. Barbier, R. Goolsby, D. Zeng. *Harnessing the crowdsourcing power of social media for disaster relief*. Tech. rep. DTIC Document, 2011 (cit. on p. 19).

[Gon+13] J. Goncalves, D. Ferreira, S. Hosio, Y. Liu, J. Rogstadius, H. Kukka, V. Kostakos. "Crowdsourcing on the spot: altruistic use of public displays, feasibility, performance, and behaviours." In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM. 2013, pp. 753–762 (cit. on p. 19).

[HT06] M. Hassenzahl, N. Tractinsky. "User experience-a research agenda." In: *Behaviour & information technology* 25.2 (2006), pp. 91–97 (cit. on p. 20).

[Har16] T. S.G.T. F. Harvard Boston Children's Hospital. *Flu Near You Community Health Project*. 2016. URL: https://flunearyou.org/ (visited on 09/12/2016) (cit. on p. 21).

[How06] J. Howe. "The rise of crowdsourcing." In: *Wired magazine* 14.6 (2006), pp. 1–4 (cit. on p. 19).

[IDC15] IDC. *Smartphone OS Market Share, 2015 Q2*. 2015. URL: http://www.idc.com/prodserv/smartphone-os-market-share.jsp (visited on 09/12/2016) (cit. on p. 41).

[Inc15] G. Inc. *Google Flu Trends*. 2015. URL: https://www.google.org/flutrends/about/ (visited on 09/12/2016) (cit. on p. 21).

[Inc16a] G. Inc. *Android Platform Versions*. 2016. URL: https://developer.android.com/about/dashboards/index.html (visited on 09/12/2016) (cit. on p. 41).

[Inc16b] G. Inc. *Cross Media Panel*. 2016. URL: https://www.crossmediapanel.com/ (visited on 09/12/2016) (cit. on p. 20).

[Inc16c] G. Inc. *Google Fit Application*. 2016. URL: https://www.google.com/fit/ (visited on 09/12/2016) (cit. on p. 55).

[Inc16d] G. Inc. *Google Forms*. 2016. URL: https://www.google.com/forms/about/ (visited on 09/12/2016) (cit. on p. 32).

[Inc16e] G. Inc. *Google Play Store*. 2016. URL: https://play.google.com/store?hl=en (visited on 09/12/2016) (cit. on p. 20).

[Inf16] Influenzanet. *Influenzanet Netherlands (All Seasons) - Participants*. 2016. URL: https://www.influenzanet.eu/en/results/ (visited on 09/12/2016) (cit. on p. 34).

[JBE07] A. Joki, J. A. Burke, D Estrin. "Campaignr: a framework for participatory data collection on mobile phones." In: (2007) (cit. on p. 20).

[JO98] D. Jobber, D. O'Reilly. "Industrial mail surveys: A methodological update." In: *Industrial Marketing Management* 27.2 (1998), pp. 95–107 (cit. on p. 27).

[Kha+11] F. Khatib, F. DiMaio, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska, I. Pichova, J. Thompson, Z. Popović, et al. "Crystal structure of a monomeric retroviral protease solved by protein folding game players." In: *Nature structural & molecular biology* 18.10 (2011), pp. 1175–1177 (cit. on p. 24).

[Kic16] Kickstarter. *Kickstarter funding statistics*. 2016. URL: https://www.kickstarter.com/help/stats (visited on 09/12/2016) (cit. on p. 19).

[Kle08] J. Kleinberg. "The convergence of social and technological networks." In: *Communications of the ACM* 51.11 (2008), pp. 66–72 (cit. on p. 26).

[Kuk+13] H. Kukka, H. Oja, V. Kostakos, J. Gonjalves, T. Ojala. "What makes you click: exploring visual signals to entice interaction on public displays." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2013, pp. 1699–1708 (cit. on p. 20).

[LC13] J. Li, C. Cardie. "Early stage influenza detection from twitter." In: *arXiv preprint arXiv:1309.7340* (2013) (cit. on p. 21).

[LES06] C. Lampe, N. Ellison, C. Steinfield. "A Face (book) in the crowd: Social searching vs. social browsing." In: *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*. ACM. 2006, pp. 167–170 (cit. on p. 26).

[Law01] E. J. Lawless. *Women escaping violence: Empowerment through narrative*. University of Missouri Press, 2001 (cit. on p. 25).

[Lon+05] I. M. Longini, A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. Cummings, M. E. Halloran. "Containing pandemic influenza at the source." In: *Science* 309.5737 (2005), pp. 1083–1087 (cit. on p. 14).

[MW10] W. Mason, D. J. Watts. "Financial incentives and the performance of crowds." In: *ACM SigKDD Explorations Newsletter* 11.2 (2010), pp. 100–108 (cit. on p. 19).

[Mag03] S. F. Magruder. "Evaluation of over-the-counter pharmaceutical sales as a possible early warning indicator of human disease." In: *Johns Hopkins University APL Technical Digest* 24 (2003), pp. 349–353 (cit. on p. 21).

[Man+10] C. Mancini, Y. Rogers, A. K. Bandara, T. Coe, L. Jedrzejczyk, A. N. Joinson, B. A. Price, K. Thomas, B. Nuseibeh. "Contravision: exploring users' reactions to futuristic technology." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2010, pp. 153–162 (cit. on pp. 15, 45).

[Mau+11] M. Mauri, P. Cipresso, A. Balgera, M. Villamira, G. Riva. "Why is Facebook so successful? Psychophysiological measures describe a core flow state while using Facebook." In: *Cyberpsychology, Behavior, and Social Networking* 14.12 (2011), pp. 723–731 (cit. on p. 26).

[Mol14] E. Mollick. "The dynamics of crowdfunding: An exploratory study." In: *Journal of business venturing* 29.1 (2014), pp. 1–16 (cit. on p. 19).

[Org+14] W. H. Organization et al. *Influenza (Seasonal) Fact Sheet No. 211, March 2014*. 2014 (cit. on p. 14).

[PA03] I. P. Pavlov, G. V. Anrep. *Conditioned reflexes*. Courier Corporation, 2003 (cit. on p. 18).

[PA10] J. Pruitt, T. Adlin. *The persona lifecycle: keeping people in mind throughout product design*. Morgan Kaufmann, 2010 (cit. on p. 25).

[Pao+14] D Paolotti, A Carnahan, V Colizza, K Eames, J Edmunds, G Gomes, C Koppeschaar, M Rehn, R Smallenburg, C Turbelin, et al. "Web-based participatory surveillance of infectious diseases: the Influenzanet participatory surveillance experience." In: *Clinical Microbiology and Infection* 20.1 (2014), pp. 17–21 (cit. on p. 21).

[Pla16] Placed. *Panel App*. 2016. URL: https://www.placed.com/panel-app (visited on 09/12/2016) (cit. on p. 20).

[QB10] W. Quesenbery, K. Brooks. *Storytelling for user experience: Crafting stories for better design*. Rosenfeld Media, 2010 (cit. on p. 25).

[RD00] R. M. Ryan, E. L. Deci. "Intrinsic and extrinsic motivations: Classic definitions and new directions." In: *Contemporary educational psychology* 25.1 (2000), pp. 54–67 (cit. on p. 17).

[Shi09] K. Shilton. “Four billion little brothers?: Privacy, mobile phones, and ubiquitous data collection.” In: *Communications of the ACM* 52.11 (2009), pp. 48–53 (cit. on p. 20).

[Sin07] J. P. Sinisi. *System and method for mobile data collection*. US Patent 7,313,759. 2007 (cit. on p. 20).

[Sta16] Statista. *Statistics and facts about Social Networks*. 2016. URL: https://www.statista.com/topics/1164/social-networks/ (visited on 09/12/2016) (cit. on p. 26).

[Sti+11] M. Stickdorn, J. Schneider, K. Andrews, A. Lawrence. *This is service design thinking: Basics, tools, cases*. Wiley Hoboken, NJ, 2011 (cit. on p. 25).

[Sud85] S. Sudman. “Mail surveys of reluctant professionals.” In: *Evaluation Review* 9.3 (1985), pp. 349–360 (cit. on p. 17).

[Swe15] R. Sweetland. *Elements of Story or Fiction*. 2015. URL: http://www.homeofbob.com/literature/genre/fiction/ficElmnts.html (visited on 09/12/2016) (cit. on p. 25).

[Tum+10] A. Tumasjan, T. O. Sprenger, P. G. Sandner, I. M. Welpe. “Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment.” In: (2010) (cit. on p. 26).

[VAD04] L. Von Ahn, L. Dabbish. “Labeling images with a computer game.” In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2004, pp. 319–326 (cit. on p. 24).

[VJ11] J. VanGeest, T. P. Johnson. “Surveying nurses: Identifying strategies to improve participation.” In: *Evaluation & the Health Professions* (2011), p. 0163278711399572 (cit. on p. 17).

[VJW07] J. B. VanGeest, T. P. Johnson, V. L. Welch. “Methodologies for improving response rates in surveys of physicians a systematic review.” In: *Evaluation & the Health Professions* 30.4 (2007), pp. 303–321 (cit. on pp. 14, 17).

[Var95] H. R. Varian. “Economic Mechanism Design for Computerized Agents.” In: *USENIX workshop on Electronic Commerce*. New York, NY. 1995, pp. 13–21 (cit. on p. 18).

[Whi09] P. Whitla. “Crowdsourcing and its application in marketing activities.” In: *Contemporary Management Research* 5.1 (2009) (cit. on p. 19).

[Yan+12] D. Yang, G. Xue, X. Fang, J. Tang. “Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing.” In: *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM. 2012, pp. 173–184 (cit. on p. 19).

All links were last followed on August 17, 2016.

**Declaration**

I hereby declare that the work presented in this thesis is entirely my own and that I did not use any other sources and references than the listed ones. I have marked all direct or indirect statements from other sources contained therein as quotations. Neither this work nor significant parts of it were part of another examination procedure. I have not published this work in whole or in part before. The electronic copy is consistent with all submitted copies.

place, date, signature