Software Design Principles for Digital Behavior Change Interventions

Lessons Learned from the MOPO Study

Lauri Tuovinen¹, Riikka Ahola², Maarit Kangas², Raija Korpelainen³,4,5, Pekka Siirtola¹, Tim Luoto⁶, Riitta Pyky²,4,5, Juha Röning¹ and Timo Jämsä²,3

¹Department of Computer Science and Engineering, University of Oulu, Oulu, Finland
²Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland
³Medical Research Center Oulu, Oulu University Hospital and University of Oulu, Oulu, Finland
⁴Centre for Life Course Health Research, University of Oulu, Oulu, Finland
⁵Department of Sports and Exercise Medicine, Oulu Deaconess Institute, Oulu, Finland
⁶Department of Cultural Anthropology, University of Oulu, Oulu, Finland

Keywords: Health Intervention, Behavior Change, Web Application, Software Design, Data Collection.

Abstract: Using the Internet as a delivery channel has become a popular approach to conducting health promotion interventions, and the evidence indicates that such interventions can be effective. In this paper we propose a set of design principles and a generic architectural model based on experiences accumulated while developing a Web-based application for a physical activation intervention. The proposed principles address the development of an intervention application as an abstract entity, a platform for gathering data for the needs of three principal stakeholder groups. The principles are derived from the purposes for which the data is gathered and the constraints that may limit the availability of desired data; by observing these principles, developers of intervention applications can identify the design trade-offs they need to make to ensure that all stakeholder needs are adequately fulfilled. An evolutionary development process is proposed as a way of gradually working toward an application that induces the desired effect on the behavior of the users.

1 INTRODUCTION

Computer software, and especially Web applications, can be convenient and effective tools for the execution of intervention studies that promote health behavior change. However, the task of translating the objectives of the intervention into software requirements is far from trivial. The standard software engineering practice of eliciting requirements from the customer is not appropriate in this context, because there is no customer in the conventional sense.

In the MOPO study¹ (Ahola et al., 2013), we have developed and deployed a gamified Web portal intended to persuade young men to improve their health behavior, especially their physical activity. In this paper we review the design challenges we needed to overcome and the approach by which we chose to address them. The paper focuses on problems that we perceive as universal to all Internet-based behavior change interventions, and on generalizing our experiences as guidelines for developers of intervention applications. The principal contribution of the paper is a conceptual framework for designing the core architecture of an intervention application and identifying factors to be taken into account in the functional specification of the application.

In particular, we observe that when considered in isolation of intervention objectives, a software application developed for a health promotion intervention is essentially a platform for collecting data from various sources and delivering it to various stakeholders. With all the details specific to a given intervention thus abstracted away, specifying such an application is reduced to identifying the information needs of the stakeholders and the sources from which the required information can be acquired. On the other hand, it is equally important to identify the constraints that may prevent the fulfillment of some of those needs. If there

---

¹http://www.tuunaamopo.fi/sivu/fl/mopo-study_in_english/
is some factor limiting the amount of data that it is feasible to collect, then it may be necessary to balance what a stakeholder would ideally expect against what is achievable in practice. Therefore, in order to discover useful design guidelines, we will need to examine both the expectations and the limitations.

Regarding the scope of the paper, our intent is not to claim that the abstraction of intervention applications as data collection platforms yields a comprehensive view, or that the list of stakeholders identified in the paper is exhaustive. What we do argue is that the abstraction provides one useful view that can and should be supplemented with views addressing other aspects and stakeholders of intervention software. The result is not a complete methodology or a universal architecture, but rather an exploration of what can be learned about the design of intervention software by adopting this particular perspective.

Section 2 of this paper presents an overview of the MOPO study and the Web application developed for it, along with a review of related work. Section 3 examines the expectations of different stakeholders and how they affect the design of intervention applications. Section 4 considers major trade-offs that may need to be made in order to design a feasible solution. Section 5 discusses the findings and some directions for future work, and Section 6 concludes the paper.

2 BACKGROUND

The medical research community has been aware of the potential of the Internet as a communications channel in public health promotion interventions since the late 1990s (Cassell et al., 1998). A vast number of Internet-based interventions have been carried out since then, targeting a variety of unhealthy behavioral patterns such as physical inactivity [e.g., (Napolitano et al., 2003; Van Den Berg et al., 2006)], smoking [e.g., (Graham et al., 2007; Saul et al., 2007)], heavy drinking [e.g., (Linke et al., 2007; Cunningham et al., 2009)], and high fat intake [e.g., (Frenn et al., 2005; Oenema et al., 2008)]. Several reviews and meta-analyses have also been conducted, indicating that Internet-based interventions have significant effects and providing recommendations for future research and practice [e.g., (Bennett and Glasgow, 2009; Tate et al., 2009; Webb et al., 2010; Brouwer et al., 2011)].

In the MOPO study, the principal target was to persuade conscription-aged Finnish men to increase their physical activity by means of a gamified Web application. The effects of the application were studied in a 6-month randomized controlled trial (Ahola et al., 2013). The study participants were recruited at the annual conscription call-ups of the Finnish Defence Forces, which every male citizen of Finland is required by law to attend in the year of his 18th birthday. Following two pilot studies carried out in 2011 and 2012 (Jauho et al., 2015), a full-scale intervention was launched in 2013, with 248 participants in the intervention group and 244 in the control group.

The interventions were carried out before the participants began their national service, so that the study tracked their activity as they went about their regular daily lives. To measure their initial physical fitness, a series of fitness tests were performed for each participant and the results were recorded. At the end of the study period, the same measurements were performed again to see if there had been any significant change in the fitness of the participants. To collect physical activity data from the study participants, each participant was loaned a Polar Active activity monitor (Kinnunen et al., 2012) and asked to wear it for the duration of the study; blinded devices, displaying only date and time, were issued to the control group.

The application developed for the study is a Web portal that provides the users with information on topics related to health, well-being and fitness, generates visualizations and verbal feedback on their physical activity, allows them to play a location-based digital game that rewards activity and mobility, and provides them with links to other relevant services and resources. To collect the activity data recorded by the study participants into a database, the participants were instructed to periodically upload the data logged by the activity monitor to the intervention server. Manual logging of exercises was also possible, to account for situations where the wearable monitor was unavailable or did not generate useful data. In addition to the activity data, information on how the application was used and perceived was collected in the study. The application itself was designed to log a number of events and statistics, and the study participants were also requested to voice their opinions on what they liked and disliked about the application.

As the number of Internet-based health promotion interventions such as the MOPO study continues to increase, we argue that there is a need for a development methodology that addresses the challenges specific to intervention applications. There is some previous work in this area that should be noted. Some authors have created primarily descriptive conceptualizations and critical reviews of various aspects of the field (Strecher, 2007; Barak et al., 2009; Danaher and Seeley, 2009; Klasnja and Pratt, 2009), while others have presented case reports of specific tools and systems (Koskinen and Salminen, 2007; Alahäivälä, 2007).
et al., 2013). Most pertinent to our work are process models, frameworks and design principles for the development of persuasive systems for health behavior change (Skinner et al., 2006; LaMendola and Krysik, 2008; Mohr et al., 2014; Nguyen et al., 2015), along with more generic design methodologies considered suitable for persuasive technology, such as participatory design (Kensing and Blomberg, 1998).

Perhaps the single most relevant piece of previous research is the BIT model (Mohr et al., 2014), which addresses the software design of intervention applications among other considerations. However, the model concentrates on a single perspective, namely building a system capable of effecting health behavior change, and this is generally true of the related work cited above. Based on what we learned from the MOPO study, we argue that the development of intervention software should be approached from multiple perspectives to account for the fact that designing an application for a health intervention is an optimization problem where maximizing the efficacy of the intervention is but one of several objectives. Another recent and relevant study is (Nguyen et al., 2015), but its topic is design of health information portals, whereas our work concerns applications that are more interactive and have more advanced application logic. On the other hand, compared to more universal design principles of persuasive technology or Web applications, our work is more attuned to the problems and requirements specific to software for behavior change interventions. There is likely to be some overlap, but we argue that intervention software is enough of a special case to merit special treatment.

3 DESIGN PERSPECTIVES

Having established that the purpose of an intervention application, when viewed on a high level of abstraction, is to collect, store and provide access to data considered important for the success of the intervention, we need to ask what exactly, then, is important. The answer depends on which perspective the application is viewed from; we discern three perspectives that should be considered:

- From the perspective of the study participants, the primary concern is operation of the application: what its features are and how they are presented to and used by the participants. The application must be able to collect and store all data required for it to deliver its functionality such that the application is relevant and interesting to the users.

- From the perspective of the researchers, the primary concern is evaluation of the application: how it performs with respect to the objectives set for the intervention. The application must be able to collect and store all data required for the researchers to track relevant attributes of the study participants and assess the performance of the application objectively and reliably.

- From the perspective of the developers, the primary concern is modification of the application: how it could be improved to make it more effective as an intervention tool. The application must be able to collect and store all data required for the developers to clearly identify directions for further development.

Although these are technically three distinct perspectives, in practice there is likely to be a substantial amount of overlap. In the MOPO study, the physical activity data was used by both the researchers, to enable them to see whether the interventions had any effect, and the study participants, in the form of feedback intended to motivate them to achieve and retain a healthy activity level. Similarly, the application usage data was used by both the researchers, to enable them to assess the extent to which the effects of the interventions were caused by the software, and the developers, to enable them to diagnose technical problems and identify potential improvements.

Among the stakeholder groups identified above, the researchers are in a key position when the functionality of the application is specified. This is because all the requirements of the application are ultimately dependent on the objectives of the study, and it is up to the researchers to hypothesize on how to achieve the objectives by means of software (operation), design a method for testing the hypothesis (evaluation), and draw conclusions on how to proceed after the study is finished (modification). Translating this knowledge into software specifications requires close cooperation between the researchers and the developers, preferably with at least some members of the development team having a dual role.

Probably the most elusive of the three design perspectives is the perspective of the study participants. Defining requirements to represent this perspective involves a considerable amount of speculation on how a given feature of the application would affect the behavior of the participants, and therefore a considerable amount of uncertainty. Consequently, this task is substantially more difficult than specifying features required for the application to support evaluation and interpretation of the study results. In the MOPO study, representatives of the target demographic participated in small-scale live testing of various development versions of the intervention application, but even with the feedback from these tests, it was diffi-
cult to predict how the participants of the intervention would ultimately respond to the application.

The challenging nature of developing software capable of inducing behavior change emphasizes the importance of the developer perspective. With so much uncertainty concerning the impact of the software, there is a high probability that the desired effect will not be achieved on the first attempt. This can be responded to by adopting an evolutionary approach where data collected during the study is used to make decisions on how the software should be revised to improve its impact on the behavior of the study participants. The development process can thus be represented as a cycle of 4 Ds, as shown in Figure 1.

Achieving behavior change in the study participants is, of course, the primary goal in an intervention study, but this does not mean that the other two perspectives are of secondary importance: the impact of the study cannot be reliably evaluated unless the necessary data is available to the researchers, and the evolutionary development process cannot advance beyond the initial cycle unless the necessary data is available to the developers. All three perspectives are thus equally indispensable, and the needs of the corresponding stakeholders are the fundamental forces that shape the intervention application when considered as a data collection platform. However, there are practical limits to what data, and how much data, it is feasible to collect in any given study; the factors that limit data collection are discussed in the next section.

4 CONSTRAINTS AND COMPROMISES

Analyzing the needs of the stakeholder groups yields a specification that states what data should ideally be collected in order to maximize the likelihood that all the identified needs can be fulfilled. In reality, however, it may be necessary to settle for a suboptimal but acceptable solution. We discern three major factors that may force such trade-offs to be made:

- **Cost**: the resources available for data collection and storage are limited by the budget of the study.
- **Privacy**: the study participants may not be willing to disclose some data that would be desirable to have for the purposes of the study.
- **Convenience**: the study participants may not be willing to participate in data collection that requires them to sacrifice their personal resources.

The cost of the resources required for data collection and storage affects both which kinds of data are available for collection and how much of them can be collected during the study. If sensor devices are to be used to collect data, then it is generally the case that the more sophisticated the desired data is, the more costly the required hardware will be. Data storage, while less likely to prove prohibitively expensive, nevertheless also needs some resources allocated to it. Furthermore, both quantity and quality of data may be constrained by personnel costs.

The privacy issue arises because some potential participants will object to being monitored, even when given the assurance that their data will be used only for the purposes of the study and kept strictly separate from any information that might be used to identify an individual participant. The more sensitive the data that the participants are requested to give access to, the more difficult it will be to find willing participants, and the fewer participants there are, the smaller the resulting body of data. Information concerning the health of an individual is generally regarded as sensitive data, making privacy a potential constraint in any study where some form of health information is collected from the participants.

Convenience is, in a sense, even more difficult to factor in than privacy, because its effect on data collection cannot be known until when the study is already underway. Even a relatively small investment of time or effort may prove enough of an inconvenience to cause some study participants to drop out of the study, at which point it may be too late to replace them. Using data collection methods that depend on the participants actively providing data therefore runs the risk of losing participants over the course of the study such that it is not easy to predict how many will stay committed. The usability of the main application is also an important factor, since the participants are likely to have a low tolerance to poor usability in an application they use mainly because they are expected to, not because they have a genuine use for it.

The MOPO study provides examples of compromises made with respect to each of the constraints.
identified above. For instance, we wanted to keep track of how certain key application usage statistics changed over the course of the study. Instead of simply sampling these numbers at regular intervals, we took periodic snapshots of the application database to account for additional research ideas that might occur during or after the collection of the dataset. However, to reduce the need for storage resources, the most voluminous component of the database — time series of the study participants’ activity levels, sampled at 30-second intervals — was left out of the snapshots. This was judged an acceptable trade-off, given that including this data in the snapshots would have multiplied their size by two orders of magnitude.

An example of a trade-off between data collection and privacy is the collection of data representing the application usage behavior of the study participants. From a technical point of view, it would have been relatively simple to log the navigation history of each participant individually, but there is a significant risk that this would have been viewed as overly invasive. As a compromise, only statistical usage logs were maintained, enabling the researchers and developers to see, for example, how much activity there had been in each section of the application. An exception was made in the case of user logins; these were logged individually, since this information was necessary for the purpose of identifying and possibly contacting study participants who had not used the application, and also useful for troubleshooting purposes.

The convenience of data collection is the area where it proved the most difficult to find a good trade-off. Uploading activity data from Polar Active to the application server required the study participants to log in to a standalone client application that communicates with the monitor through Flowlink, a small peripheral plugged into the USB port of the computer running the client. To avoid overrunning the memory buffer of Polar Active, each study participant needed to repeat the upload process at least once every 21 days. In theory, this would not seem to be an excessive inconvenience, but in practice, it was necessary to send out reminders to the participants and to offer them chances to win prizes by uploading their activity data. It was also found that any glitches in the upload process, or in any other application function, could have a substantial demoralizing effect, causing some participants to stop trying rather than contact the developers for a solution (Luoto et al., 2014). The precarious motivation of the participants poses a considerable problem, since it would be important for the developers to get their feedback on how to make the application more captivating, but lacking the motive, the participants may not even use the application enough to be able to give any meaningful feedback.

With regard to each constraining factor, it is imperative to make a deliberate decision on a trade-off, to ensure that the outcome is both acceptable and achievable. Overemphasizing either of these puts the success of the study at risk: if too much weight is placed on achievability, there is a risk of diluting the objectives of the study, whereas if acceptability is weighted too heavily, there is a risk that the study will fail to collect enough data to support the objectives. Either way, the further off balance the trade-off is, the more difficult it will be to achieve good results.

5 DISCUSSION

Table 1 shows some metrics based on login statistics that can be used to evaluate the acceptance of our application in the 2012 and 2013 interventions. The most notable change is the increase in the percentage of the intervention group who tried the application at least once; there is no improvement when we look at the ratio of the number of users who logged in multiple times to the number of all users. Thus it seems that although the 2013 study had more success in persuading participants to start using the application, the users lost interest in the application at a similar rate, with only a relatively small number of them becoming regulars. This does not mean that the studies did not yield any positive results, but data from other sources suggests that these were mostly caused by other motivating factors such as the activity monitor.

Some of the difficulty of motivating the participants of the MOPO study can probably be attributed to the target population being particularly challenging to address effectively. The adverse health effects of sedentary lifestyle generally do not occur until later in life, so it is difficult to get young people to feel that

| Table 1: Some metrics computed from application login statistics for the 2012 and 2013 interventions. Application users are defined as intervention group members who logged in at least once, returning users as those who logged in at least twice, and monthly users as those who logged in at least four times between the beginning of the intervention in September and the end of the year. |
|-----------------|-------|-------|
| **Intervention group (N)** |
| Application users (Nu) | 141   | 248   |
| Ratio of Nu to N | 0.47  | 0.65  |
| Returning users (Nr)  | 49    | 117   |
| Ratio of Nr to Nu   | 0.74  | 0.73  |
| Monthly users (Nm)   | 24    | 59    |
| Ratio of Nm to Nu    | 0.36  | 0.37  |
participating in the intervention is important to them personally, if their only motivation is the abstract knowledge that they have an elevated risk of suffering from certain medical conditions in the future. Under these circumstances, when executing a study based on monitoring the participants while they act freely in an uncontrolled environment, it may prove a considerable challenge to ensure that the participants do not forget or neglect their commitment.

Related to this is our observation that it is not advisable to rely on study participants to report problems encountered while using the application. People who use an application because they need it for something have an incentive to report glitches, namely the hope that they will be patched in a future update, but this logic fails in the case of intervention applications, because the users may have no particular interest in using the application. In this case their willingness to report problems again depends on the depth of their commitment, and an uncommitted study participant, when encountering a problem that effectively prevents them from using the application, may simply give up rather than seek help to solve the problem. Lesser usability issues, such as an unintuitive or unattractive user interface, are similarly more discouraging if the user is not internally motivated to continue using the application.

This uncertainty concerning the commitment of the study participants makes it all the more important to design the application to automatically gather data for development purposes. Automatic data collection cannot completely substitute human feedback, especially on subjective issues such as how interesting or persuasive the application is, but it can be sufficient for the developers to diagnose situations where the application performs inadequately. Most importantly, automatic notification, whenever possible, is the fastest way to inform the developers when a failure occurs.

We can summarize the design imperatives derived from our experience as follows:

- Designers of intervention applications should identify the constraints that may prevent the application from collecting desired data, and design trade-offs to ensure that the outcome is satisfactory to all stakeholders. In particular, the data collection objectives should be set such that:
  - achieving them is possible using the resources available for the execution of the intervention.
  - achieving them is possible without unduly limiting the privacy of the study participants.
  - achieving them is possible without unduly inconveniencing the study participants.

Although these principles are all technically derived from a single study, the fact that there were two pilots before the full-scale intervention means that we had multiple opportunities to accumulate experience on the design and execution of Internet-based health promotion interventions. This was particularly useful in that it enabled us to observe the importance of collecting data to support the specification of the next version of the intervention application, both automatically and by eliciting input from members of the target population. Thus we came to apply the evolutionary development approach outlined in Section 3, and as a result of this, both the first pilot study and the second one led us to identify major software modifications that should be made before the next iteration. They also improved our understanding of what data should be collected; for example, the full metabolic equivalent (MET) data from Polar Active was not used in the first pilot, and some additional digested values calculated from the MET data during upload were used only in the final intervention.

Given the amount of development work done in the MOPO study, along with the fact that we have abstracted away all application-specific details, we would argue that the design principles proposed here are generalizable to other Internet-based interventions. Furthermore, the architecture of the MOPO software is also potentially generalizable to a wider range of applications. If we apply the abstraction used in this paper to the architecture, we can reduce it to a data repository and a set of interfaces through which the three stakeholder groups access the data:

- The data upload interface, by which whatever data on the study participants is being monitored in the intervention is sent to the application.
- The user interface, by which the study participants submit requests to the application and have information presented to them.
- The administrator interface, by which the researchers and developers have unlimited access to the data stored by the application.
To benefit from this abstract architecture, it could be implemented in the form of a software framework that would provide a generic foundation on which applications for many different digital behavior change interventions could be built. The architecture of an intervention application would then look like Figure 2, where the data interface APIs are provided by the framework and the interface modules, which implement the application-specific functionality, are coded by the application developers, using data import and export services provided by the APIs. The interface APIs handle communication with the application database, either directly or via the intermediate analytics API, which provides services for generating various digested presentations from the original data.

The data repository is divided into two conceptually distinct sections: static content, which only the administrators can write to, and collected data, which holds the data acquired over the course of the study by monitoring the participants. The internal architecture of the repository should reflect the modularity of the data interfaces; due to the evolutionary nature of the development process, data collection requirements are likely to change and the core architecture must be able to accommodate these changes. With several different stakeholder needs to be fulfilled, and several factors constraining the quality or quantity of the data that can be acquired to fulfill them, it is especially important to ensure that the design of the application does not introduce any additional constraints. Making the data interfaces and storage structures modular will eliminate or at least significantly reduce the risk of not being able to use an important data source because the developers were not aware of it when the application was originally specified.

6 CONCLUSION

Web applications hold considerable potential as tools for executing effective health interventions promoting behavior change. In this paper we proposed a set of design principles for such applications based on the observation that an intervention application can be viewed as an abstraction whose principal functionality consists of gathering data from various sources and presenting it to stakeholders in various formats. From this observation, we derived three principles that prescribe identification of stakeholder needs, and another three that prescribe identification of constraints. Furthermore, we concluded that an evolutionary development approach is suitable for tackling the problem of predicting how the users of the application will respond to it. These findings are the result of hands-on experience of developing and deploying three iterations of a Web-based application designed to persuade young men to change their physical activity behavior.

ACKNOWLEDGEMENTS

The MOPO study was supported by the Finnish Ministry of Education and Culture; Juho Vainio Foundation; Centre for Military Medicine Finland; Northern Ostrobothnia Hospital District; Centre for Economic Development, Transport and the Environment of North Ostrobothnia; European Social Fund; Finnish Funding Agency for Technology and Innovation; and European Regional Development Fund.

REFERENCES


