Process-Driven Data Collection with Smart Mobile Devices

Johannes Schobel, Marc Schickler, Rüdiger Pryss, and Manfred Reichert

Institute of Databases and Information Systems, Ulm University, Germany
{johannes.schobel, marc.schickler, ruediger.pryss, manfred.reichert}@uni-ulm.de

Abstract. Paper-based questionnaires are often used for collecting data in application domains like healthcare, psychology or education. Such paper-based approach, however, results in a massive workload for processing and analyzing the collected data. In order to relieve domain experts from these manual tasks, we propose a process-driven approach for implementing as well as running respective mobile business applications. In particular, the logic of a questionnaire is described in terms of an explicit process model. Based on this process model, in turn, multiple questionnaire instances may be created and enacted by a process engine. For this purpose, we present a generic architecture and demonstrate the development of electronic questionnaires in the context of scientific studies. Further, we discuss the major challenges and lessons learned. In this context the presented process-driven approach offers promising perspectives in respect to the development of mobile data collection applications.

1 Introduction

During the last years smart mobile applications have been increasingly used in business environment. Examples include applications for task management and location-based services. In particular, smart mobile devices offer promising perspectives in respect to mobile data collection as well [13]. For example, data could be collected with sensors (e.g., pulse sensor), communicating with the smart mobile device [23] or with form-based end-user applications [12]. Such a mobile data collection becomes necessary, for example, in the context of clinical trials, psychological studies, and quality management surveys.

In order to enable mobile data collection, specific knowledge on how to implement such smart mobile applications is required on one hand. On the other, domain-specific knowledge is needed, which is usually not available to application developers. Consequently, costly interactions between domain experts and IT experts are required. To reduce overall efforts, a framework for rapidly developing and evolving mobile data collection applications shall be developed. In particular, respective business applications shall be easy to maintain for non-computer experts as well. Our overall vision is to enable domain experts to develop mobile data collection applications themselves at a high level of abstraction. Specifically, this paper focuses on the process-driven design, implementation and enactment of mobile questionnaires that support end users with their daily data collection tasks.

As application domain for demonstrating the benefits of our approach we choose psychological studies. Here, domain experts mostly use paper-based questionnaires for collecting required data from subjects. However, such a paper-based data collection shows several drawbacks, e.g., regarding the structuring and layout of a questionnaire (e.g., questions may still be answered even if they are no longer relevant) as well as the later analysis of answers (e.g., errors might occur when transferring the paper-based collected data to electronic worksheets).

To cope with these issues and to understand the differences between paper-based and electronic questionnaires in a mobile context, first of all, we implemented selected questionnaire applications for smart mobile devices and applied them in real world application settings [3, 21, 6]. In particular, we were able to demonstrate that mobile electronic questionnaires relieve end users from costly manual tasks, like the transfer, transformation and analysis of the collected data. As a major drawback, however, all these applications were hard-coded and their implementation required considerable interactions with end users. As a consequence, the respective applications were neither easy to maintain nor extensible. In order to overcome the gap between the domain-specific design of a questionnaire and its technical implementation enacted on smart mobile devices, therefore, an easy to handle, flexible and generic *questionnaire system* is required.

From the insights we gained during the practical use of the aforementioned mobile applications as well as from lessons learned in related implemented projects [20], we elicited the requirements for electronic questionnaire applications that allow for a flexible mobile data collection. In order to evaluate whether the use of process management technology contributes to the satisfaction of these requirements, we mapped the logic of a complex questionnaire from psychology to a process model, which was then deployed to a process engine. In particular, the process model served as basis for driving the execution of questionnaire instances at runtime. Note that this mapping allows overcoming many of the problems known from paper-based questionnaires. In turn, the use of a process modeling component and a process execution engine in the given context raised additional challenges. Especially, the implemented questionnaire runs on a mobile device and communicates with a remote process engine to enact psychological questionnaires. As a major lesson, we learned that process management technology may not only be applied in the context of business process automation, but also provides a promising approach for generating mobile data collection applications. In particular, a process-driven approach enables end users (i.e., non-computer experts) to develop mobile electronic questionnaires as well as to deploy them on smart mobile devices.

The contributions of this paper are as follows:

- We discuss fundamental problems of paper-based questionnaires and present requirements regarding their transfer to smart mobile devices.

- We provide a mental model for mapping the logic of questionnaires to process models and illustrate this mental model through a real-world application scenario from psychology.
- We present a generic architecture for data collection applications on smart mobile devices. This architecture can be applied to model, visualize and enact electronic questionnaires. In particular, it uses process models to define and control the flow logic of a questionnaire.
- We share fundamental insights we gathered during the process of implementing and evaluating mobile data collection applications.

The remainder of this paper is organized as follows: Section 2 discusses issues related to paper-based questionnaires. Further, it elicitates the requirements that emerge when transferring a paper-based questionnaire to an electronic version running on smart mobile devices. Section 3 describes the mental model we suggest for meeting these requirements. In Section 4, we present the basic architecture of an approach for developing mobile data collection applications. Section 5 provides a detailed discussion, while Section 6 presents related work. Finally, Section 7 concludes the paper with a summary and outlook.

2 Case Study

In a case study involving 10 domain experts, we analyzed more than 15 paper-based questionnaires from different domains, including questionnaires used in the context of psychological studies. Our goal was to understand issues that emerge when transferring paper-based questionnaires to smart mobile applications. Section 2.1 discusses basic issues related to *paper & pencil* questionnaires, while Section 2.2 elicitates fundamental requirements for their electronic mobile support.

2.1 Paper-Based Questionnaires

We analyzed 15 paper-based questionnaires from psychology and medicine. In this context, a variety of issues emerged. First, in the considered domains, a questionnaire must be *valid*. This means that it should have already been applied in several studies, and statistical evaluations have proven that the results obtained from the collected data are representive. In addition, the questions are usually presented in a neutral way in order to not affect or influence the subject (e.g., patient). Creating a valid instrument is one of the main goals when setting up a psychological questionnaire. In particular, reproducible and conclusive results must be guaranteed. Furthermore, a questionnaire may be used in two different modes. In the *interview mode*, the subject is interviewed by a supervisor who also fills in the questionnaire; i.e., the supervisor controls which questions he is going to ask or skip. This mode usually requires a lot of experience since the interviewer must also deal with questions that might be critical for the subject. The other mode we consider is *self-rating*. In this mode, the questionnaire is handed out to the subject who then answers the respective questions herself; i.e., no supervision is provided in this mode

Another challenging issue of paper-based questionnaires concerns the *analysis* of the data collected. Gathered answers need to be transfered to electronic worksheets, which constitutes a time-consuming and error-prone task. For example, when filling in a questionnaire, typographical errors or wrong interpretations of given answers might occur. In general, both sources of error (i.e., errors occurring during the interviews or transcription) decrease the quality of the data collected, which further underlines the need of an electronic support for data collection.

In numerous interviews we conducted with 10 domain experts from psychology, additional issues have emerged. Psychological studies are often performed in developing countries, e.g., surveying of child soldiers in rural areas in Africa [3]. *Political restrictions* regarding data collection further require attention and influence the way in which interviews and assessments may be performed by domain experts (i.e., psychologists). Since in many geographic regions the available infrastructure is not well developed, the data collected is usually digitalized in the home country of the scientists responsible for the study. Taking these issues into account, it is not surprising that psychological studies last from *several weeks up to several months*. From a practical point of view, this raises the problem of allocating enough space in luggage to transfer the paper-based questionnaires safely to the home country of the respective researcher.

Apart from these *logistic problems*, we revealed issues related to the interview procedure itself. In particular, it has turned out that questionnaires often need to be *adapted by authorized domain experts to a particular application context* (e.g., changing the language of a questionnaire or adding / deleting selected questions). In turn, these adaptations must be propagated to all other interviewers and smart mobile devices respectively to keep the results valid and comparable.

Considering these issues, we had additional discussions with domain experts, which revealed several requirements as discussed in the next section.

2.2 Requirements

In the following, we discuss basic requirements for the mobile support of electronic questionnaires. We derived these requirements in the context of case studies, literature analyses, expert interviews, and hands-on experiences regarding the implementation of mobile data collection applications [21, 6]. Especially, when interviewing domain experts, fundamental requirements could be elicitated. The same applies to the paper & pencil questionnaires we analyzed.

Basic requirements derived from the interviews are listed below:

R1 (Mobility). The process of collecting data should be flexible and usually requires extensive interactions. Data may have to be collected even though no computer is available at the place the questionnaire should be filled in. For example, consider data collection at the bedside of a patient in a hospital or interviews conducted by psychologists in a meeting room. Computers are often disturbing in such situations, particularly if the interviewer is "hiding" himself behind a screen. To enable flexible data collection, the device needs to be portable instead. Further, it should not distract the participating actors in communicating and interacting with each other.

R2 (Multi-User Support). Since different users may interact with a mobile questionnaire, multi-user support is crucial. Furthermore, it must be possible to distinguish between different user roles (e.g., interviewers and subjects) involved in the processing of an electronic questionnaire. Note, that a user may possess different roles. For example, he could be interviewer in the context of a specific questionnaire, but subject in the context of another one.

R3 (Support of Different Questionnaire Modes). Generally, a questionnaire may be used in two different modes: interview and self-rating mode (cf. Section 2.1). These two modes of questioning diverge in the way the questions are posed, the possible answers that may be given, the order in which the questions are answered, and the additional features provided (e.g., freetext notes). In general, mobile electronic questionnaire applications should allow for both modes. Note, that this requirement is correlated with R2 as the considered roles determine the modes available for a questionnaire.

R4 (Multi-Language Support). The contents of a questionnaire (e.g., questions and field labels) may have to be displayed in different languages (e.g., when conducting a psychological study globally). The actor accessing the questionnaire should be allowed to choose the preferred language.

R5 (Maintainability). Questionnaires evolve over time and hence may have to be changed occasionally. Therefore, it should be possible to quickly and easily change the structure and content of an electronic questionnaire; e.g., to add a question, to edit the text of a question, to delete a question, or to change the order of questions. In particular, no programming skills should be required in this context; i.e., domain experts (e.g., psychologists) should be able to introduce respective changes at a high level of abstraction.

Due to the lack of space, not all requirements are listed here. More requirements with respect to the user interface and native application design are summarized in [22]. Especially, requirement R5 constitutes a major challenge. It necessitates a high level of abstraction when defining and changing electronic questionnaires, which may then be enacted on various smart mobile devices. To cope with this challenge, we designed a specific mental model for electronic questionnaires, which will be presented in Section 3.

3 Mental Model

To transfer paper-based questionnaires into electronic ones and to meet the requirements discussed, we designed a mental model for the support of mobile electronic questionnaires (cf. Figure 1). According to this model, the logic of a paper-based questionnaire is described in terms of a process model, which is then deployed to a process management system. The latter allows creating and executing process (i.e., questionnaire) instances.

Generally, a process model serves as template for specifying, enacting and evolving structured processes based on process management systems. In addition,

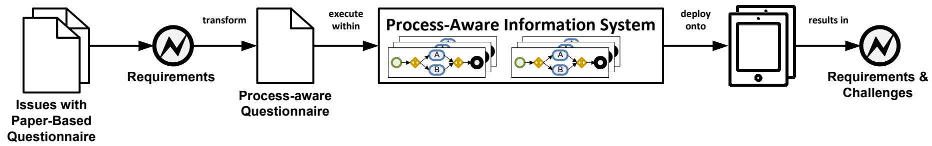

Fig. 1. Mental model

adaptive process management systems allow for dynamic changes of process instances in order to cope with unpredictable situations [19]. In the following, we show that process models and process management technology are not only useful in the context of business process automation, but may be applied for mobile data collection as well. However, this raises additional challenges (cf. Section 5). This paper will show how to realize a process-aware questionnaire system whose models guide the users in filling in electronic questionnaires.

3.1 Process Model and Process Instances

As opposed to traditional information systems, *process-aware information systems* (PAIS) separate process logic from application code. This is accomplished based on process models, which provide the schemes for executing the respective processes [25]. In addition, a process model allows for a visual (i.e., graph-based) representation of the respective process, comprising the process steps (i.e., activities) as well as the relations (i.e., control and data flow) between them. For control flow modeling, both control edges and gateways (e.g., ANDsplit, XORsplit) are provided.

A process model P is represented as a directed, structured graph, which consists of a set of nodes N (of different types NT) and directed edges E (of different types ET) between them. We assume that a process model has exactly one start node ($NT = StartFlow$) and one end node ($NT = EndFlow$). Further, a process model must be connected; i.e., each node n can be reached from the start node. In turn, from any node n of a process model, the end node can be reached. In this paper, we solely consider block-structured process models. Each branching (e.g. parallel or alternative branching) has exactly one entry and one exit node. Furthermore, such blocks may be nested, but are not allowed to overlap [17]. In turn, data elements D correspond to global variables, which are connected with activities through data flow edges ($ET\_DataFlow$). These data elements can either be read ($ReadAccess$) or written ($WriteAccess$) by activities of the process model [16]. Figure 3 shows an example of a process model.

In turn, a process instance I represents a concrete case that is executed according to a process model P. In general, for a given process model multiple instances may be created and concurrently executed. Thereby, the state of a particular instance is defined by the markings (i.e., states) of its nodes and edges as well as the values of its data elements. Altogether, respective information corresponds to the execution history of an instance. Usually, a process engine relies on a set of execution rules describing the constraints for which a particular

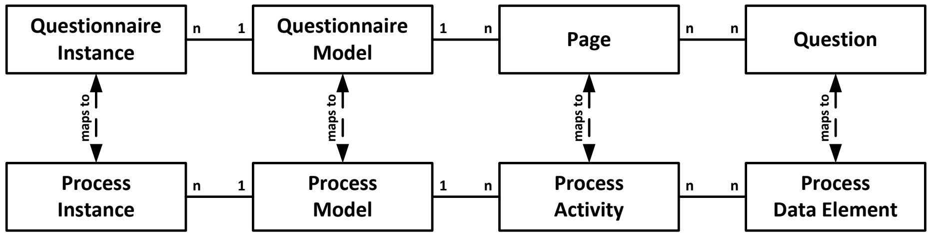

Fig. 2. Mapping a Questionnaire Model to a Process Model

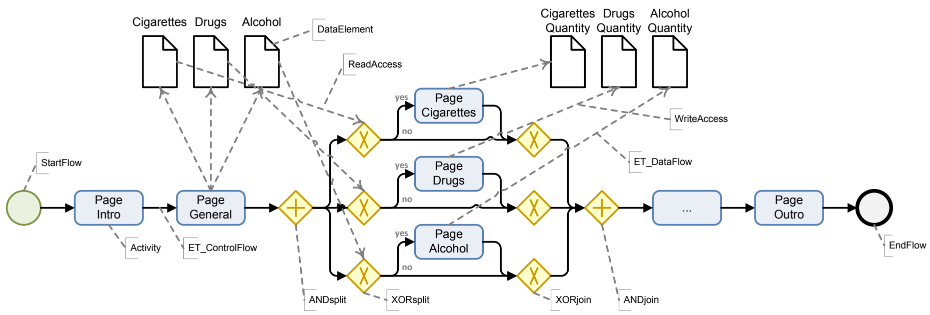

Fig. 3. Application Scenario: an abbreviated Questionnaire with Annotations

activity may be activated [16]. If the end node of the process model is reached, the respective process instance terminates.

3.2 Mapping a Questionnaire to a Process Model

Our mental model for enabling a process-driven enactment of questionnaires is as follows: We capture both the logic and the content of a questionnaire in a corresponding process model. Accordingly, pages of the questionnaire logically correspond to process activities, whereas the flow between these activities specifies the control flow logic of the questionnaire. The questions themselves are mapped to process data elements, which are connected with the respective activity. There are separate elements containing the text of a question, which can be read by the activity. Moreover, there are data elements that can be written by the activity. The latter are used to store the given answers for a specific question. Figure 2 presents an overview of the mapping of the elements of a questionnaire to the ones of a process model.

To illustrate the process-driven modeling of electronic questionnaires, we present a scenario from psychology. Consider the process-centric questionnaire model from Figure 3 whose logic is described in terms of BPMN 2.0 (Business Process Model and Notation). To establish the link between process and questionnaire model, we annotate the depicted graph with additional labels.

The processing of the questionnaire starts with the execution of activity *Page Intro*, which presents an introductory text to the participant interacting with the electronic questionnaire. This introduction includes, for example, instructions on how to fill in the questionnaire or interact with the smart mobile device. After completing this first step in the processing of the questionnaire, activity *Page General* becomes enabled. In this form-based activity, data elements *Cigarettes*, *Drugs* and *Alcohol* are written. More precisely, the values of these data elements correspond to the answers given for the questions displayed on the respective page of the questionnaire. For example, the question corresponding to data element *Cigarettes* is as follows: *"Do you smoke?"* (with the possible answers *"yes / no"*). After completing activity *Page General*, an *AND gateway* (ANDsplit) becomes enabled. In turn, all outgoing paths of this ANDsplit (i.e., parallel split node) become enabled and are then executed concurrently. In the given application scenario, each of these paths contains an *XOR gateway* (XORsplit), which reads one of the aforementioned data elements to make a choice among its outgoing paths. For example, assume that in *Page General* the participant has answered question *"Do you smoke?"* with *"yes"*. Then, in the respective XORsplit, the upper path (labeled with *"yes"*) will be chosen, which consists of exactly one activity, i.e., *Page Cigarettes*. In the context of this activity, additional questions regarding the consumption of cigarettes will be displayed to the actor. This activity and page, respectively, is exemplarily displayed in Figure 4. Assume further that question *"Do you take drugs? (yes / no)"* has been answered with *"no"* in *Page General*. Then, activity *Page Drugs* will be skipped as the lower path (labeled with *"no"*) of the respective XOR split is chosen. As soon as all three branches are completed, the *ANDjoin* will become enabled and the succeeding activity be displayed. We omit descriptions of the other activities of the questionnaire model due to lack of space. The processing of a questionnaire ends with activity *Page Outro*. Note that, in general, an arbitrary number of questionnaire instances processed by different participants may be created.

Figure 4 gives an impression of the *Page Cigarettes* activity. It displays additional questions regarding the consumption of cigarettes. This page is layouted automatically by the electronic questionnaire application based on the specified process model, which includes the pages to be displayed (cf. Figure 3). Note that the user interface is created based on the data elements which contain the text of the questions as well as the possible answers to be displayed (i.e., the answers among which the user may choose).

3.3 Requirements for Process-Based Questionnaires

When using process management technology to coordinate the collection of data with smart mobile devices, additional challenges emerge. In particular, these are related to the modeling of a questionnaire as well as the process-driven execution of questionnaire instances on smart mobile devices.

Since questionnaire-based interviews are often interactive, the participating roles (e.g., interviewer and interviewed subject) should be properly assisted when interacting with the smart mobile device. For example, it should be possible for them to start or abort questionnaire instances. In the context of long-running questionnaire instances, in addition, it might be required to interrupt an interview

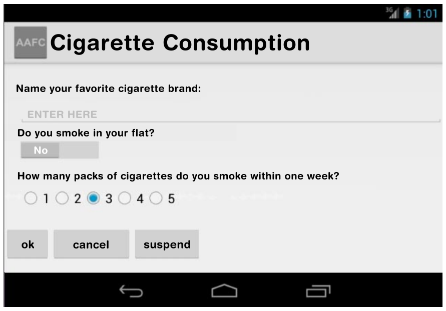

Fig. 4. Activity "Page Cigarettes"

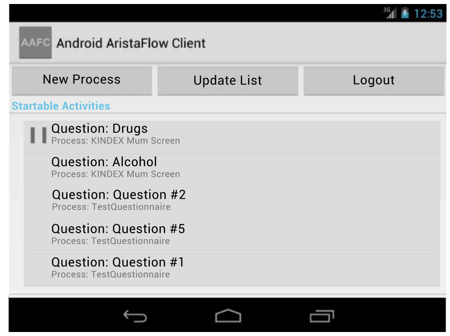

Fig. 5. Startable Activities for an Actor

and continue it later. For this purpose, it must be possible to suspend the execution of a questionnaire instance and to resume it at a later point in time (with access to all data and answers collected so far). In the context of long-running interviews, one must be able to display an entire questionnaire and process model respectively. Therefore, already answered questions should be displayed differently (e.g., different color) compared to upcoming questions. Note that this is crucial for providing a quick overview about the current progress.

Since domain experts might not be familiar with existing process modeling notations like BPMN 2.0, an easy-to-understand, self-explaining and domain-specific process notation is needed. In addition, the roles participating in a questionnaire should be provided with specific views on the process model (i.e., questionnaire), e.g., hiding information not required for this role [8]. For example, a subject might not be allowed to view subsequent questions in order to ensure credibility of the given answers.

Regarding the execution of the activities of a questionnaire (i.e., pages) additional challenges emerge.

The questions of a (psychological) questionnaire might have to be answered by different actors each of them possessing a specific role. For example, follow-up questions related to the subject involved in a psychological questionnaire have to be answered by a psychologist and not by the subject itself. In order to avoid bad quality of the data collected, actors should be further assisted when interacting with the smart mobile device; e.g., through error messages, help texts, or on-the-fly validations of entered data. Consequently, the electronic questionnaire application must ensure that only those questions are displayed to an actor that are actually intended her. Figure 5 shows the startable activities, currently available for a specific actor using the smart mobile device.

To foster the subsequent analysis of the data collected, the latter needs to be archived in a central repository. Furthermore, additional information (e.g., the time it took the subject to answer a particular question) should be recorded in order to increase the expressiveness of the data collected. Finally, anonymization of the data might have to be ensured as questionnaires often collect personal data

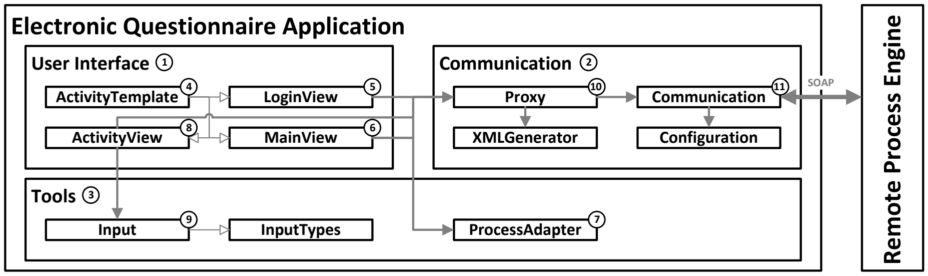

Fig. 6. Architecture of the Electronic Questionnaire Application

and privacy constitutes a crucial issue. In certain cases, it might also become necessary to dismiss the results of an already answered question.

Taking these general requirements into account, we designed an architecture for an electronic questionnaire application.

4 Architecture and Implementation

This section introduces the architecture we developed for realizing electronic questionnaires. In particular, the latter run on smart mobile devices and interact with a remote process engine. This architecture is presented in Section 4.1. Since this paper focuses on the requirements, challenges and lessons learned when applying state-of-the-art process management technology to realize electronic questionnaires, we will not describe the architecture of the remote process management system in detail (see [4, 18] for respective work). The general architecture of our electronic questionnaire application is depicted in Figure 6.

4.1 Electronic Questionnaire Application

The electronic questionnaire application is divided into three main packages, which are related to the *user interface* ①, the *communication* ② with the external process engine, and useful *tools* for interacting with the client ③.

The user interface representing a particular page of the questionnaire is represented by an *ActivityTemplate* ④, which provides basic methods for the questionnaire (e.g., to start or stop an activity). In turn, the *LoginView* ⑤ is used to query the user credential and to select an available role for this actor (e.g., $name = JohnDoe, role = Interviewer$). Furthermore, the *MainView* ⑥ provides a list (e.g., worklist) with the pages currently available for the user interacting with the questionnaire. The list items are represented using the *ProcessAdapter* ⑦. Since the user interface of a questionnaire is generated dynamically depending on the underlying process model that has been deployed on the process engine, a user interface generator is needed. This service is provided by the *ActivityView* ⑧. To interact with the device, different classes of the *Input* ⑨ elements used within a questionnaire are provided. These classes provide the necessary logic to interact with the input elements as well as the corresponding graphical representation of this element. As certain input elements are platform-specific (e.g., there is no spinning wheel for standard desktop applications), missing input elements might be rendered differently, depending on the underlying platform (e.g., the spinning wheel on iOS could be rendered as a dropdown element on a normal computer).

The entire communication with the external process engine should be handled by a *Proxy* ⑩ service. The latter is capable of generating the necessary request messages, which are then converted to SOAP request messages by the *Communication* ⑪ service and sent to the process engine. The response messages (e.g., the next page to display) sent by the process engine are then received by the *Communication* and decomposed by the *Proxy*. Afterwards, the data within this message is visualized in the *ActivityView*, which includes the already mentioned user interface generator as well.

4.2 Proof-of-Concept Prototype

To validate the feasibility of the described architecture as well as to apply it in a real setting, we implemented a proof-of-concept prototype for Android. The prototype application was then used to verify the prescribed mental model, and to detail the requirements regarding the execution of process-aware questionnaires. Furthermore, additional insights into the practical use of this application by domain experts in the context of their studies could be gathered. We were able to meet the requirements presented in Section 2.2 when implementing the questionnaire client application, even though certain drawbacks still exist. In order to enable domain experts, who usually have no programming skills, to create a mobile electronic questionnaire, we implemented a fully automated user interface generator for the mobile application itself. In addition, we were able to provide common types for questions used within a questionnaire (e.g., likert-scale, free-text or yes-no-switches). These types are automatically mapped to appropriate input elements visualized within the application.

5 Discussion

This section discusses our approach and reflects on the experiences we gained when applying state-of-the-art process management technology to support mobile data collection with electronic questionnaires. Since we applied an implemented questionnaire in a psychological study, we were also able to gain insights into practical issues as well.

The presented approach has focused on the development of mobile business applications enabling flexible data collection rather than on the design of a development framework. Therefore, we have used an existing process modeling editor for defining the logic of electronic questionnaires. However, since the domain experts who have been using our questionnaire application have been unfamiliar with the BPMN 2.0 modeling notation, a number of training sessions were required. Afterwards, they were able to create their own questionnaires.

In particular, the abstraction introduced by the use of process models for specifying questionnaire logic was well understood by domain experts. However, the training sessions have shown that there is a need for a more user-friendly, domain-specific modeling notation, enabling domain experts to define questionnaires on their own. In particular, such a domain-specific modeling language needs to be self-explaining and easy to use. Further, it should hide modeling elements not required in the given use case. Note that BPMN 2.0 provides many elements not needed for defining the logic of electronic questionnaires. Consider, for example, the AND gateways, which allow modeling parallel execution paths. Regarding the use case of mobile data collection, it does not matter which path is going to be evaluated first. In addition, the elements for modeling a questionnaire should have a clear semantics and be expressive enough. Therefore, an activity should be represented as page-symbol to add context-aware information to the questionnaire model.

As we further learned in our case study, the creation and maintenance of a questionnaire constitutes a highly interactive, flexible and iterative task. In general, the editing of already existing, but not yet published questionnaires, should be self-explaining. Basic patterns dealing with the adaptation of the logic of a questionnaire (e.g., moving a question to another position or adding a new question) should be integrated in a modeling editor to provide tool support for creating and managing questionnaires.

As discussed in Section 3.3, we use process management technology for both modeling and enacting electronic questionnaires. Accordingly, the created questionnaire model needs to be deployed on a process engine. Regarding the described client server architecture (cf. Section 4.1), all process (i.e., questionnaire) models are stored and executed on the server running the process engine. Keeping in mind that mobile questionnaires might be also used in areas without stable Internet connection, any approach requiring a permanent internet connection between the mobile client and the process engine running on an external server will not be accepted. In order to cope with this issue, a light-weight process engine is required, which can run on the smart mobile device. We have started working in this direction as well (e.g., see [14, 15]).

Since the user interface of the electronic questionnaire is automatically generated based on the provided process model, the possibilities to customize the layout of a questionnaire are rather limited. From the feedback we had obtained from domain experts, however, it became clear that an expressive layout component is needed that allows controlling the visual appearance of a questionnaire running on smart mobile devices. Among others, different text styles (e.g., bold), spacing between input elements (e.g., bottom spacing), and absolute positioning of elements should be considered. In addition, the need for integrating images has been expressed several times.

Since we use process-driven electronic questionnaires for collecting data with smart mobile devices, the answers provided by the actors filling in the questionnaire could be directly transferred to the server. This will relieve the actors from time-consuming manual tasks. Furthermore, as there exists a process model describing the flow logic of the questionnaire as well as comprehensive instance data (e.g., instance execution history), process mining techniques for analyzing questionnaire instances might be applied [1]. In addition, Business Intelligence Systems [2] could reveal further interesting aspects with respect to the data collected in order to increase the expressiveness of the analysis. Such systems would allow for a faster evaluation and relieve domain experts from manual tasks such as transferring the collected data into electronic worksheets.

Finally, we have experienced a strong acceptance among all participating actors (e.g., interviewers, domain experts, and subjects) regarding the practical benefits of electronic questionnaire applications on smart mobile devices. Amongst others this was reflected by a much higher willingness to fill out an electronic questionnaire compared to the respective paper-based version [21, 6]. Furthermore, a higher motivation to complete the questionnaire truthfully could be observed. Of course, this acceptance partly results from the modern and intuitive way to interact with smart mobile devices.

6 Related Work

There exists a variety of questionnaire systems available on the market. In general, these systems can be classified into two groups: *online services* [24] and *standalone applications* [5]. Due to the fact that a questionnaire might contain sensitive information (e.g., the mental status of a subject or personal details), online surveys are often not appropriate for this type of data collection applications. As another limitation of online systems, local authorities do often not allow third-party software systems to store the information of a subject. However, these applications also must deal with privacy issues. Standalone applications usually offer possibilities to create a questionnaire, but do not deal with the requirements discussed in this paper. Furthermore, they lack a flexible and mobile data collection. Usually, respective questionnaires are displayed as web applications, which cannot be used when no Internet connection is available.

To the best of our knowledge, the process-aware enactment of questionnaires on smart mobile devices has not been considered comprehensively by other groups so far. In previous studies, we identified crucial issues regarding the implementation of psychological questionnaires on smart mobile devices [3, 6, 21]. In these studies, we aimed at preserving the validity of psychological instruments, which is a crucial point when replacing paper-based questionnaires with electronic ones. Although the implemented applications have already shown several advantages in respect to data collection and analysis, they have not been fully suitable for realizing psychological questionnaires in the large scale yet. In particular, maintenance efforts for domain experts and other actors were considerably high. More precisely, changes of an implemented questionnaire (or its structure) still had to be accomplished by computer experts, due to its implementation. Therefore, we aim to integrate a process-aware information system to overcome this limitation.

Focusing on the complete lifecycle of paper-based questionnaires and supporting every phase with mobile technology has actually not been considered by other groups so far. However, there exists related work regarding mobile data collection. In particular, mobile process management systems, as described in [13, 9], could be used to realize electronic questionnaires. However, this use case has not been considered by existing mobile process engines so far.

The QuON platform [11] is a web-based survey system, which provides a variety of different input types for the questions of a questionnaire. As opposed to our approach, QuON does not apply a model-based representation for specifying a questionnaire. Another limitation of QuON is its restriction to web-based questionnaires. Especially, in psychological field studies, the latter will result in problems as the QuON platform does not use responsive webdesign.

Movilitas applies SAP Sybase Unwired Platform to enable a mobile data collection for business scenarios [10]. In turn, the Sybase Unwired Platform constitutes a highly adaptive implementation framework for mobile applications, which interacts with a backend, that provides all relevant business data. Further research is required to show whether this approach can be applied to realize mobile electronic questionnaires in domains like psychology or healthcare as well.

Finally, [7] presents a set of patterns for expressing user interface logic based on the same notation as used for process modeling. In particular, a transformation method applies these patterns to automatically derive user interfaces based on a bidirectional mapping between process model and user interface.

7 Summary and Outlook

In this paper, limitations of paper-based questionnaires for data collection were discussed. To deal with these limitations, we derived characteristic requirements for electronic questionnaire applications. In order to meet these requirements, we suggested the use of process management technology. According to the mental model introduced, a questionnaire and its logic can be described in terms of a process model at a higher level of abstraction. To evaluate our approach, a sophisticated application scenario from the psychological domain was considered. We have shown how a questionnaire can be mapped to a process model.

In the interviews we conducted with domain experts as well as from other mobile business applications we implemented, general requirements for flexible mobile data collection on smart mobile devices were elaborated. These cover major aspects like a secure and encrypted communication. Note that the latter is crucial, especially in domains like medicine and psychology, which both deal with sensitive information of the subjects involved. We further presented an architecture enabling mobile data collection applications based on a smart mobile device and a process engine. As another contribution, we demonstrated the feasibility of our proof-of-concept application. Several features as well as problems regarding the implementation and communication with the server component, hosting the process engine, have been highlighted. Finally, we discussed the benefits of using process-driven questionnaires for mobile data collection.

In future work, we plan to extend our approach with additional features. Currently, we are working on a mobile process engine running on the smart mobile device. In turn, this will enable a process-driven enactment of questionnaire instances, even if no Internet connection is available. We consider this as a fundamental feature for enabling flexible data collection applications on smart mobile devices. A first prototype is already implemented and integrated into the presented electronic questionnaire application. However, this will cause new problems, as the questionnaire models must be properly synchronized among multiple devices (e.g., if changes were made to the questionnaire model). In addition, we are working on a questionnaire modeling notation. This is crucial to allow domain experts, who are unfamiliar with standard process modeling notations. This domain-specific questionnaire modeling notation shall be easy to understand, but still precise enough for defining process-aware questionnaires. The notation will be part of a *generic questionnaire system* supporting the complete lifecycle of a questionnaire. Furthermore, we started to integrate sensors measuring vital signs in order to gather other information about subjects during interviews [23]. As a major benefit of the framework, we expect higher data quality, shorter evaluation cycles and a significant decrease of workloads.

References

1. van der Aalst, W.M., Reijers, H.A., Weijters, A.J., van Dongen, B.F., Alves de Medeiros, A., Song, M., Verbeek, H.: Business process mining: An industrial application. Information Systems 32(5), 713–732 (2007)
2. Anandarajan, M., Anandarajan, A., Srinivasan, C.A.: Business intelligence techniques: a perspective from accounting and finance. Springer (2003)
3. Crombach, A., Nandi, C., Bambonye, M., Liebrecht, M., Pryss, R., Reichert, M., Elbert, T., Weierstall, R.: Screening for mental disorders in post-conflict regions using computer apps - a feasibility study from burundi. In: XIII Congress of European Society of Traumatic Stress Studies (ESTSS) (2013)
4. Dadam, P., Reichert, M.: The ADEPT Project: A Decade of Research and Development for Robust and Flexible Process Support - Challenges and Achievements. Computer Science - Research and Development 23(2), 81–97 (2009)
5. Electric Paper Evaluationssysteme: EvaSys. http://www.evasys.de/ (November 2013), last visited: 06. July 2014
6. Isele, D., Ruf-Leuschner, M., Pryss, R., Schauer, M., Reichert, M., Schobel, J., Schindler, A., Elbert, T.: Detecting adverse childhood experiences with a little help from tablet computers. In: XIII Congress of European Society of Traumatic Stress Studies (ESTSS) Conference (2013)
7. Kolb, J., Hübner, P., Reichert, M.: Automatically Generating and Updating User Interface Components in Process-Aware Information Systems. In: 20th Int'l Conf. on Coop. Inf. Sys. pp. 444–454. No. 7565 in LNCS, Springer (2012)
8. Kolb, J., Reichert, M.: A flexible approach for abstracting and personalizing large business process models. Applied Computing Review 13(1), 6–17 (2013)
9. Kunze, C.P., Zaplata, S., Lamersdorf, W.: Mobile processes: Enhancing cooperation in distributed mobile environments. Journal of Computers 2(1), 1–11 (2007)
10. Movilitas: Movilitas Consulting AG. http://www.movilitas.com/ (November 2013), last visited: 06. July 2014
11. Paul, D., Wallis, M., Henskens, F., Nolan, K.: QuON: A Generic Platform for the Collation and Sharing of Web Survey Data. International Conf. on Web Inf. Sys. and Techn. (2013)

12. Pryss, R., Langer, D., Reichert, M., Hallerbach, A.: Mobile Task Management for Medical Ward Rounds - The MEDo Approach. In: 1st Int'l Workshop on Adaptive Case Mgmt., BPM'12 Workshops. pp. 43–54. No. 132 in LNBIP, Springer (2012)
13. Pryss, R., Musiol, S., Reichert, M.: Collaboration Support Through Mobile Processes and Entailment Constraints. In: 9th IEEE Int'l Conference on Collab. Comp.: Networking, Applications and Worksharing. IEEE Computer Society Press (2013)
14. Pryss, R., Tiedeken, J., Kreher, U., Reichert, M.: Towards flexible process support on mobile devices. In: Proc CAiSE'10 Forum - Information Systems Evolution. pp. 150–165. No. 72 in LNBIP, Springer (2010)
15. Pryss, R., Tiedeken, J., Reichert, M.: Managing Processes on Mobile Devices: The MARPLE Approach. In: CAiSE'10 Demos (2010)
16. Reichert, M., Dadam, P.: ADEPTflex-Supporting Dynamic Changes of Workflows Without Losing Control. Journal of Intelligent Information Systems, Special Issue on Workflow Management Systems 10(2), 93–129 (1998)
17. Reichert, M., Dadam, P.: Enabling Adaptive Process-aware Information Systems with ADEPT2. In: Handbook of Research on Business Process Modeling, pp. 173–203 (2009)
18. Reichert, M., Dadam, P., Rinderle-Ma, S., Jurisch, M., Kreher, U., Goeser, K.: Architecural Principles and Components of Adaptive Process Management Technology. In: PRIMIUM - Process Innovation for Enterprise Software, pp. 81–97. No. P-151 in Lecture Notes in Informatics (LNI) (2009)
19. Reichert, M., Weber, B.: Enabling Flexibility in Process-Aware Information Systems: Challenges, Methods, Technologies. Springer, Berlin-Heidelberg (2012)
20. Robecke, A., Pryss, R., Reichert, M.: DBIScholar: An iPhone Application for Performing Citation Analyses. In: CAiSE Forum-2011. Proc CAiSE'11 Forum, vol. Vol-73. CEUR Workshop Proceedings (2011)
21. Ruf-Leuschner, M., Pryss, R., Liebrecht, M., Schobel, J., Spyridou, A., Reichert, M., Schauer, M.: Preventing further trauma: KINDEX mum screen - assessing and reacting towards psychosocial risk factors in pregnant women with the help of smartphone technologies. In: XIII Congress of European Society of Traumatic Stress Studies (ESTSS) Conference (2013)
22. Schobel, J., Schickler, M., Pryss, R., Maier, F., Reichert, M.: Towards process-driven mobile data collection applications: Requirements, challenges, lessons learned. In: 10th Int'l Conf. Web Inf. Sys. and Techn. (WEBIST 2014). pp. 371–382 (2014)
23. Schobel, J., Schickler, M., Pryss, R., Nienhaus, H., Reichert, M.: Using Vital Sensors in Mobile Healthcare Business Applications: Challenges, Examples, Lessons Learned. In: 9th Int'l Conf. Web Inf. Sys. and Techn. (WEBIST 2013). pp. 509–518 (2013)
24. SurveyMonkey: SurveyMonkey: Free online survey software & questionnaire tool. http://www.surveymonkey.com/ (May 2013), last visited: 06. July 2014
25. Weber, B., Reichert, M., Mendling, J., Reijers, H.: Refactoring large process model repositories. Computers in Industry 62(5), 467–486 (2011)