D3.2 User Studies Plan

Mads Schaarup Andersen(ALX), Laura Lynggaard Nielsen(ALX)

The project SODA has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731583.

Project Information

Scalable Oblivious Data Analytics

Project number:	731583
Strategic objective:	H2020-ICT-2016-1
Starting date:	2017-01-01
Ending data:	2019-12-31
Website:	https://soda-project.eu/

Document Information

Title:	Plan for user studies				
ID:	D3.2	Type:	R	Dissemination level:	PU
		Month:	M12	Release date:	December 22, 2017

Contributors, Editor & Reviewer Information

Contributors (person/partner)	Mads Schaarup Andersen (ALX)
	Laura Lynggard Nielsen (ALX)
Editor (person/partner)	Mads Schaarup Andersen (ALX)
Reviewer (person/partner)	Anna Zsófia Horváth (GU)
	Meilof Veeningen (PHI)

Release History

Release number	Date issued	Release description / changes made
1.0	December 22, 2017	First release to EU

SODA Consortium

Full Name	Abbreviated Name	Country
Philips Electronics Nederland B.V.	PHI	Netherlands
Alexandra Institute	ALX	Denmark
Aarhus University	AU	Denmark
Göttingen University	GU	Germany
Eindhoven University of Technology	TUE	Netherlands

Table 1: Consortium Members

Executive Summary

This document is Deliverable 3.2 and contains a plan for the user studies to be conducted in the SODA project. We have planned to study the data subjects by looking at four different patient groups distributed among two subcontractors. We will study patients with complicated and uncomplicated diabetes at Aarhus University Hospital (AUH) and people with early and later stage dementia at Lancaster University (LU). The aim is to look at similarities and differences in attitudes towards data sharing by getting a deep understanding of data in relation to the patients' condition and how that fits into their treatment and daily life. Furthermore, we will study end users by looking at data analysists and epidemiologists at both subcontractors. This document covers the methodology of the user studies as well as a tentative plan for them.

About this Document

Role of the deliverable

This deliverable aims to provide a plan for the user studies to be carried out in the SODA project.

Relationship to other SODA deliverables

The deliverable constitutes the plan for the user studies that form the basis of the D3.3 User studies analysis. Furthermore, D6.1 Ethics approval is a prerequisite of the studies planned in this deliverable.

Relationship to other versions of this deliverable

N/A

Structure of this document

Section 1 is the introduction. Section 2 details the studies of the data subject followed by the studies of the end users in Section 3. Section 4 contains the timeline of the user studies.

1 Table of Contents

Release History 4
SODA Consortium 5
Executive Summary 6
About this Document 7
Role of the deliverable 7
Relationship to other SODA deliverables 7
Relationship to other versions of this deliverable 7
Structure of this document 7
1 Table of Contents 8
2 Introduction 9
3 Data subjects 10
3.1 Methodology 10
3.1.1 Tentative plan 11
4 End Users 12
4.1 Methodology 12
4.1.1 Tentative plan 13
5 User Studies Timeline 14

2 Introduction

This document constitutes deliverable 3.2 of the SODA project. The aim of this is to provide a plan for how the user studies will be conducted and explain the idea behind the methodology.

In relation to the SODA project, we are interested in looking at two different types of users: the patients[1] who provide their medical data for medical research and data analysis (data subjects), and the users who carry out the medical research and data analysis based on the data subjects' data (end users).

To get access to the user groups in the SODA project, a subcontractor approach has been chosen. The idea behind this is to get access to both relevant medical personnel working with data and data analysts as well as people with medical conditions and to let the subcontractors assist in facilitating workshops actively engaging the users. We have two subcontractors: Aarhus University Hospital (AUH), who will cover patients with complicated and uncomplicated diabetes; and Lancaster University (LU), who will cover patients with early and later stage dementia. Both subcontractors will in addition facilitate access to medical researchers and/or data analysts.

In this document, the studies of the data subjects will be explained, followed by the studies of the end users.

[1] Our preliminary studies suggest that though diabetes and dementia patients are considered patients by the health system, they do not consider themselves patients, but rather their condition is a premise in their life.

3 Data subjects

Regarding the data subjects, the end goal is to provide them with an understanding of what is going on with their data and produce a user experience in which they can truly be said to be in control, which among other things mean that they provide informed consent for their data to be used. On a high level it is important to get an understanding of how data subjects provide data for medical research and data analysis and how they consent to the use of this data. To understand this in depth, we need to explore: (1) how data subjects understand their data; (2) how they understand the data flows; (3) how they see their data being beneficial to themselves and/or others with the same condition; (4) how they perceive the privacy surrounding the use of their data; and, (5) on a fundamental level, what they understand to be their data. We will do this using a bottom-up approach where we start out by exploring what data data subjects are generating and using in the treatment of their condition. Once knowledge about this is established, we aim to create a new user experience tailored to the user groups in question in which they get the knowledge and control they need to make an informed decision about whether or not to share their data in a SODA based system.

The reason for choosing four different types of patient groups was outlined in the proposal. These groups were chosen under the assumption that they would have differing issues and motivations around data and data sharing.

3.1 Methodology

We will take a qualitative approach, utilizing a number of different methods. The plan sketched out below is a preliminary plan and we might change methods depending on our findings when analyzing the outcomes of activities. For example, if a workshop shows that more group discussion is needed, another workshop will be added to the plan.

The main methods that will be used are: workshops, semi-structured interviews, observations, user-centric design, and prototyping. We will use these activities to investigate the following topics:

- Explore the data subjects' relationship to the data that relates to their condition.
    - What type of data is generated?
    - What type of data is relevant?
    - How is it shared?
    - How is consent handled?
    - Where and to whom does data flow?
    - Are data subjects comfortable with the way data is handled today?
    - What are the privacy concerns, if any?
    - Are there any types of data collection and/or use that the data subjects feel could be useful, but that is not currently taking place?
- Explore how data subjects' data is currently flowing from the user to the hospitals/data analysts and/or other places.
    - How is data collected and by whom?
    - With what aims is data being collected? Who decides what data is relevant?
    - What consent mechanisms exist? (paper, oral, digital, etc.)
    - Do data subjects understand what they are consenting to?
    - How is data handled during analysis; in what form is it utilized (anonymized/full access); and who has access to them?
    - Is useful data generated and not being used in treatment and if so, why?
    - When is consent given?
- Understand the key elements in making consent informed.

    - How much understanding of the data analysis process and goals is needed for the data subjects?
    - How much understanding of the security model is needed for the data subjects?
    - When does it make sense to ask for consent?
- Understand the motivation for data subjects to share or not to share data.
    - What makes data subjects want to share or not to share certain types of data?
    - Who are data subjects willing to share their data with?
    - To which degree can a better understanding of consent and the security models help motivate sharing?
- Designing a user experience for data handling with SODA technologies that enables data subjects to better understand the value chain that their data enter into when giving consent.
    - How can we make the experience of data sharing more transparent for data subjects?

3.1.1 Tentative plan

- Workshop 1:
    - Goal: Get a basic understanding of the users' relationship to data in connection with their condition.
    - Method: Co-creative workshop with data subjects (and carers/legal guardians of data subjects in the diabetes case).
- Workshop 2:
    - Goal: Get an understanding of users' interaction with their medical contacts in terms of data.
    - Method: Co-creative workshop with data subjects (and carers/legal guardians of data subjects in the diabetes case).
- Observations:
    - Goal: Observe how data subjects interact with the medical domain and observe how consent to data usage is handled.
    - Method: Ethnographic observations.
- Qualitative Interviews:
    - Goal: Get a deeper understanding of the concepts of: data sharing, consent, value of data, motivations for sharing, and privacy concerns. In addition, to explore how the data subjects understand the security around data sharing.
    - Method: A series of in-depth semi-structured qualitative interviews with relevant data subjects (and carers/legal guardians of data subjects in the diabetes case)
- Iterative user experience design:
    - Goal: Develop a user experience in which data subjects are able to make informed decisions about providing their data to a system based on SODA technologies.
    - Method: An iterative prototyping approach in which a prototype for the consent user experience is created based on the previous results and is shown and tested with data subjects.

4 End Users

In terms of end users, the end goal is to enable them to do data analysis and medical research on larger amounts of data than would be possible without the technologies developed in the SODA project, i.e., secure multi-party computation (MPC) and differential privacy. This is important to understand because that while MPC and differential privacy provide the possibility to work on larger data sets, the technologies also impose limitations on computation speed and data visibility. To explore this, we want to get an understanding of how the analysts work with data today, including which tools they use. We will use these insights to design a new process and/or tool which will enable the larger datasets based on SODA technology and where possible provide input for the design of the demonstrators of Work Package 4.

As opposed to the data subjects, we assume that the different groups of end users will have similar issues and concerns with introducing SODA technologies. This means that the process might not necessarily be the same across the end users provided by AUH and LU.

4.1 Methodology

We will take a qualitative approach which will utilize a number of different methods. The plan sketched out below is the current plan, but, we might change methods depending on the outcomes of activities when they are analyzed. For example, if a workshop shows that more discussion in groups are needed, another workshop will be added to the plan.

The main methods that will be used are: workshops, semi-structured interviews, observations, contextual inquiry, user-centric design, and prototyping. We will use these things to investigate the following topics:

- Understand how end users do their job today.
    - What tasks, tools, and focus areas are involved in their work?
    - How do they handle data?
    - How do they use data visually, if so?
    - What is the computational cost of the analysis done?
    - How frequently are computationally costly analysis run on data?
    - What is the value chain surrounding data analysis?
    - Which actors and interests are involved in the process?
    - Who requests the analysis?
    - Who receives the results and how?
    - Are there any cooperation with other parties around the data analysis?
- Understand the limitations and possibilities of SODA technologies.
    - Which data is available for end users to directly access?
    - What is the data output of SODA technologies?
    - How does SODA technologies impact the runtime of the analysis made by end users?
    - How does SODA technologies impact the human actors surrounding the technology and analysis processes?
- Understand SODA technologies impact trust in the outcome of an analysis.
    - How does the lack of visual inspection of data impact the trust in the results, for end users as well as the people receiving the results?
        - Is the experience of making a decision based on the result different for a data analyst and a medical professional respectively?
        - What are the positive and negative impacts?
    - How do other aspects of the SODA technologies impact the level of trust in results??
- Design a new process and/or tool to enable SODA technologies in end users' work.

- How do we design around the limitations and possibilities of SODA technologies?

4.1.1 Tentative plan

- Pilot interviews:
  - Goal: Get a basic understanding of the tasks involved in doing data analysis within the medical domain, including tools, methods, etc.
  - Method: Semi-structured interviews.
- Contextual Inquiry.
  - Goal: Get a deeper understanding of the process of the end user.
  - Method: Contextual Inquiry with 10 end users.
- Workshop:
  - Goal: Get a broader understanding of how end users approach the task of doing data analysis. How do they view data and security? Do they think of the privacy of their data subjects? Is there any mechanism to verify e.g., statistical results? What is their trust in complex security protocols?
  - Method: Co-creative workshop with end users.
- Iterative prototype design:
  - Goal: Develop a new process and/or tool to facilitate data analysis for end users.
  - Method: An iterative prototyping approach based on workshops in which a new user experience and/or tool surrounding data analysis is created based on the previous results and is shown and tested with end users.

5 User Studies Timeline

The following outlines the tentative plan of the user studies. The reason that it points back in time is to capture the already conducted work.

	2017	**2018**				**2019**
	Q4	*Q1*	*Q2*	*Q3*	*Q4*	*Q1-Q4*
LU data subjects	Workshop 1	Workshops	Interviews observations	Interviews	Prototyping	
AUH data subjects		Workshops	Interviews observations	Interviews	Prototyping	
LU end users		Workshops	Contextual Inquiry	Interviews	Prototyping	
AUH end users	Pilot Interviews	Contextual Inquiry		Interviews	Prototyping	