Exploring the practicality of measuring mental models through log file analysis: An application in the health informatics domain

Iris ten Klooster

University of Twente

Master Thesis

Faculty of Behavioural, Management and Social Sciences (BMS)

Department of Cognitive Psychology and Ergonomics

August, 2018

1st supervisor: dr. M.L. Noordzij

2nd supervisor: dr. S.M. Kelders

Acknowledgments

I wrote this Master's thesis for completing my Master's in Human Factors and Engineering Psychology. A Master in which I have learned a lot of this relevant and fascinating field of work. In order to conclude this Master’s, I have been allowed to delve deeper into involving the user within the Agile process.

Acknowledgement to dr. M.L. Noordzij, for supervising and providing feedback throughout the composition of this thesis, as well as to dr. S.M. Kelders for providing feedback during the writing process. Moreover, I want to thank the software company for the possibility they offered me to analyse the log files of a large healthcare system. I would also like to thank my supervisor and colleagues of this software company for giving me insight into the complexity that comes with developing healthcare information technologies. The company will remain anonymous at the request of the software company. For this reason, the company is called Developos in this study, which is not the real name of the company that is involved. Lastly, I would like to express my gratitude to my family and friends for the support they gave me during writing this thesis.

Deventer, 2018

Table of contents

Abstract

Background: There is an increasing use of log file analysis within health informatics research. Log files give an objective view of users' behaviour in software systems, and provides valuable insights in how healthcare information technology is used. Gaining insight into users' behaviour may give indirect clues relevant to the objective of User-Centred Design (UCD), by obtaining a strong understanding of the intended user. Yet, to our knowledge, no studies investigating how log file analyses can be used in practice within this Agile process have been carried out. Therefore, the following question was raised: 'How may log file analyses contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?'. *Method:* The study consisted of an (1) observational phase, in which log files were analysed by summarizing sequential patterns, heatmapping and clustering. In the (2) qualitative phase, these results were discussed in a semi-structured interview with developers to explore for which practical purposes these results may be used. *Results:* In the observational phase of this study, it appeared that summarizing sequentisl patterns provides insight into prominent order that occurs within the user sessions, and heatmapping provides insight into the transitions between lesser-visited pages within the system. Moreover, clustering users' behaviour provides an insight in the variation that exists within user sessions. In the qualitative phase of this study, log file analyses have proved to be useful for stating hypotheses, comparing paths, benchmarking and prioritizing within the Agile development process. *Discussion:* The multiple practical applications supports the suggestion that log files analyses is useful for increasing the focus on the user within an Agile process. Future research is needed for determining the level of detail in which users' behaviour should be analysed.

Keywords: log file analysis, health information technology, Agile, user involvement, Lean UX

Exploring the practicality of measuring mental models through log file analysis: An application in the health informatics domain

There is an increasing use of log file analysis within health informatics research. Log files give an objective view of users' behaviour in software systems, and provides valuable insights in how healthcare information technology is used (Sieverink, Kelders, Poel, & Gemert-pijnen, 2017). Gaining insight into users' behaviour may give indirect clues relevant to the objective of User-Centred Design (UCD), by obtaining a strong understanding of the intended user (Interaction Design Foundation, n.d.). UCD is a vital determinant of a system' success, since, amongst others, it leads to quality improvement, resource savings, increased user satisfaction and company competitiveness (Ardito, Buono, Caivano, Costabile, Lanzilotti, Bruun, & Stage, 2011). However, an increasing number of software companies work according to the Agile development method, which is difficult to integrate with UCD practices (Ardito, Buono, Caivano, Costabile, Lanzilotti, Bruun, & Stage, 2011; Bak, Nguyen, Risgaard, & Stage, 2008). Instead of focusing on user involvement, Agile focuses on effective coding and customer input (Sohaib & Khan, 2010). Log file analyses provide the opportunity to an increased understanding of the user without actually having to involve the user. Yet, to our knowledge, no studies investigating how log file analyses can be used in practice within this Agile process have been carried out. Therefore, we use a systematic approach to log file analysis in this study, and present these to developers to explore how they can use it for increasing the focus on the intended user (UCD) within an Agile process.

Figure 1 shows a graphical representation of the focus of the Agile method (customer input and effective coding), UCD (mental models of the user) and log file analysis. Moreover, this figure shows how the designer communicates with the user through the interface of the system. In turn, this interface modifies the users' mental model of the system. The users' mental model serves as guides while navigating through the system, which is represented by the line pointing to users' behaviour. User's behaviour can be described on three distinct time-bands, to explain various mechanisms of users' behaviour. This study focuses on the transaction level, as shown by the black dotted lines in the model.

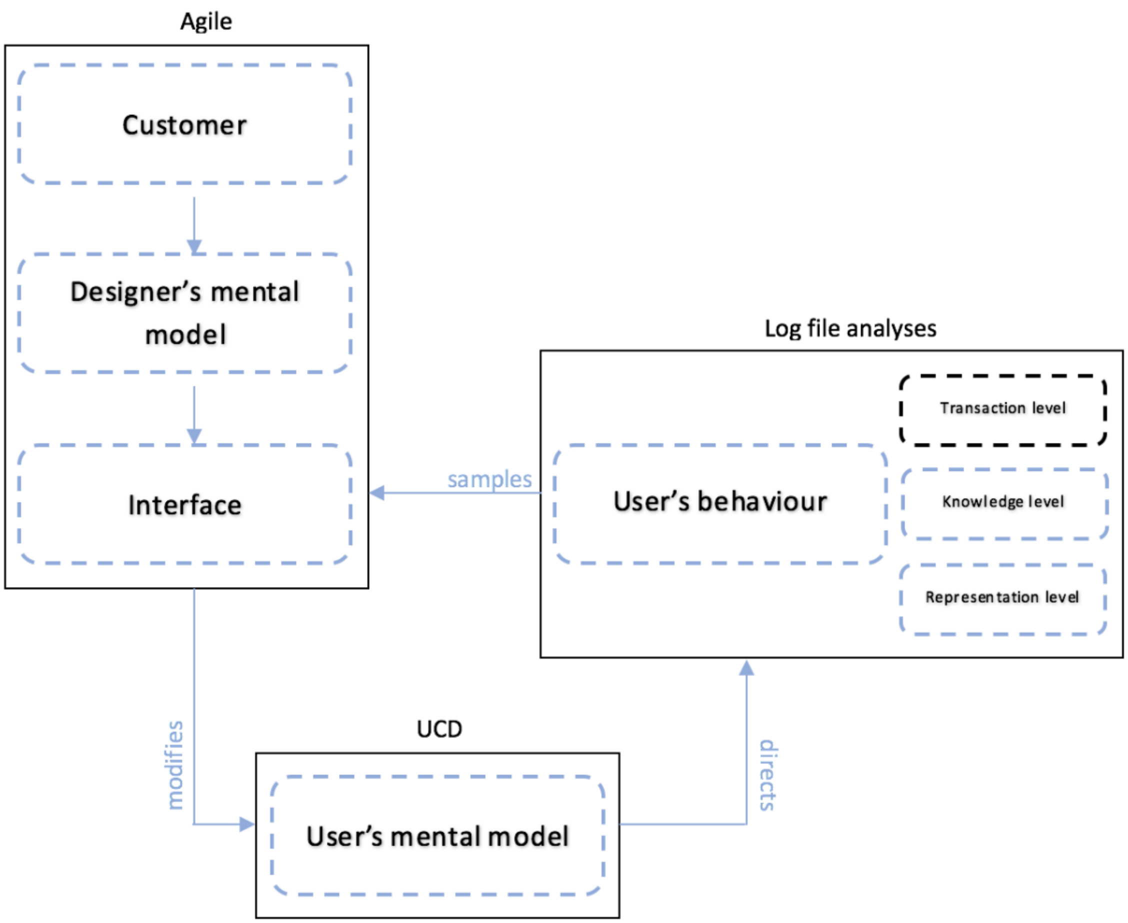

Figure 1. Conceptual model of this study, which is a is a synthesis of elements of the Designer's Model, the User's Model, and the System Image as illustrated by Norman (2013, p. 32), the Perception-action cycle as illustrated by Neisser (1976), and the time-bands as described by Newell (1989).

In the following sections, we will describe the elements of figure 1, and their relations, in more depth. First, we introduce the reader to the Agile development process and the tensions between this method and UCD practices. We then describe the relation between log file analysis and users' mental models. Lastly, we propose an approach to log file analysis that fits with describing users' behaviour on the transaction level.

Agile and User-Centred Design

Agile is defined in the Oxford Dictionary as "Able to move quickly and easily" (n.d.), and refers to multiple methods within software development, which acts as a counterpart to the previously used Waterfall method. Examples of these Agile methods within software development are Scrum, Extreme Programming, and DSDM. Whereas waterfall methods follow predefined steps in a fixed order, Agile is characterized by its small software releases with rapid iterations of two to four weeks. Every iteration can be seen as a small project in itself, and after every iteration the product is shown and tested, and the product and process are evaluated (Dybå & Dingsøyr, 2008). Since Agile consists of rapid iterations, working according to an Agile development cycle allows for adapting to changing circumstances (Dybå & Dingsøyr, 2008). the two main tensions between this agile method and UCD are explained below.

Firstly, User Centred Design (from now on UCD), which is the starting point for usability and UX, is not a priority during the Agile process (Moreno, Seffah, Capilla, Sanchez-Segura, 2013; Ardito, Buono, Caivana, Costabile, & Lanzilotti, 2014). This coincides with the differences in focus of these methods (Sohaib & Khan, 2010). Agile focuses on effective coding by means of acceptation and unit testing, while UCD is an approach to place the user at the heart of the development process (Sohaib & Khan, 2010). Working according to an Agile development method brings a so-called “developer mindset”, meaning that developers have their main interest in efficiency of code, which has no impact on end users of the product (Ardito, Buono, Caivano, Costabile, Lanzilotti, Bruun, & Stage, 2011; Bak, Nguyen, Risgaard, & Stage, 2008).

Secondly, during iterations of the Agile development cycle, developers try to constantly work together with customers by close communication (Dybå & Dingsøyr, 2008; Sohaib & Khan, 2010). However, since the customer is not necessarily the user of the system, Agile development is characterized by a focus more on the customer than the user, compared to User Centred Design (Sohaib & Khan, 2010). Although both Agile and UCD put emphasis on stakeholder involvement, which is identified as a key factor for integrating Agile and UCD (Brhel, Meth, Maedche, Werder, 2015), the customer often fulfils both the role of customer and user within the Agile development method (Brhel, Meth, Maedche, Werder, 2015). This may lead to determining wrong requirements for the system, and to setting aside the requirements of the user (Jokela & Abrahamsson, 2004).

Log files and users’ mental models

Log files are a documentation of users’ behaviour within a particular system (Dumais, Jeffries, Russel, Tang, & Teevan, 2014). These are automatically produced, and provide information about the time of the event, the url visited, and about either posting or retrieving information from the system. Given that these log files are produced while the user is in her or his natural environment, this meets the demand for more naturalistic methods (Schraagen, Militello, Ormerord, & Lipshitz, 2016). The pursuit of naturalistic methods stems from the claim that controlled experimental studies have not led to results that are generalizable to real-world situations in complex domains (Kushniruk, 2002).

According to Norman (2013), users' behaviour is strongly impacted by their mental models of the system. His framework for describing mental models consisted of four components, of which three are relevant for the purpose of this study (Norman, 2013). The system itself is captured in the component 'target system', on which an interface is created to allow the user to interact with the software (Norman, 2013). The users' mental models are the internal representations of this software, on which users make inferences on how to carry out tasks (2013). According to Nielsen (2010) users' mental models become more accurate and complete while interacting with the system.

When working according to a user-centered approach, Agile team members try to fit the software to the users' mental models. However, developers' mental models consist mainly of how the system works, while users' mental models consist of beliefs on how to use the system (Nielsen, 2010). This means that the software developed in an Agile method regularly doesn't fit with the mental models of the user (Nielsen, 2010). One way to prevent this difference between the developers' and users' mental models is to get users in front of the interface to see whether this mental model is different from the developers' mental model (Nielsen, 2010) (e.g. by carrying out the thinking- aloud method). There are a number of methods to find out whether the developers' mental models differ from the users' mental models.

These methods can be described using Newell's bands of behaviour to understand in which processes insight is obtained through these UCD methods (Newell, 1989; Russel, 2014). These bands entail three distinct time-frames, namely the 'representation level', the 'knowledge level', and the 'transaction level' (Newell, 1989; Russel, 2014). This division of levels was to explain how different mechanisms come to light along different time-frames. A higher level means that mechanisms at lower levels become invisible, while other mechanisms become visible. For example, explaining user behaviour on the representation level, explains why users look at a certain part of the interface, while on the transaction level users' behaviour is described as a sequence of events. On the transaction leven, users' behaviour within distinct parts of pages become invisible (e.g. clicking, scrolling within page), while users' behaviour related to navigating between these pages become visible (e.g. navigating from 'News-page' to the 'Search-page').

The highest level, the transaction level, describes the system as a network comprised of nodes, connected to each other through links (Schraagen, 2017). Contrary to describing users' behaviour on the lower level, analysis here focuses on navigating between pages, independent of the content within pages. Log file analysis is consistent to this level of analysing users' behaviour. Since this analysis entails a longer time-span, behaviour described on the lower levels become invisible, while different users' behaviour becomes visible.

The aforementioned time-bands all operate concurrently, and by using a number of methods which are related to all three levels, a more varied insight is obtained in users' behaviour. Integrating this way of data collection (analysing all three levels) into the health information technology, is in line with Kushniruk's argument for carrying out a wider range of methods (2002).

Log file analyses

One method that fits well with describing users' behaviour on the transaction level, is analysing users' behaviour as a sequence of events as described by Scholz (2016). This way of analysing provides insight into the order in which the user navigates through the system, which relates to a longer time span (transaction level), while the more classical approach to log file analysis only describes quantities of usage (e.g. number of times pages are visited, mean duration of visiting a page). In this analysis, there are some key concepts that play a role. Firstly, a user session is a set of page views between logging in and logging out to the system, for a particular user at one particular website. Secondly, 'nodes' are the distinct urls in the log files, grouped into page categories (e.g. 'News page'). Lastly, within user sessions, users navigate between these 'nodes'. A user navigating from page-category A to page-category B ('nodes') forms a 'link' between these page-categories.

Summarizing sequential patterns.

Summarizing sequential patterns can be done for obtaining insight into the number of times that 'nodes' (or pages) and 'links' appear within users' sessions. Moreover, it describes how many times page-categories appear at a specific step within the user session. For example, the number of times that a specific page is visited as a first step within all users' sessions. These sorts of analyses are used to describe usage of the system in the past.

Heatmapping.

Heatmapping provides functionalities for predicting 'links' between 'nodes' of users' behaviour, which is similar to what has been called 'local rules' in social network analysis (Butts, 2008). This means that these analyses centres on the properties of distinct 'nodes' or page-categories. Heatmapping means that the probability of navigating from one specific page-category to the other page-categories are calculated and these probabilities are shown in a heatmap. These provide insight in local usage behaviour, since these probabilities are centred around the distinct page-categories. The probabilities are calculated by means of markov chain modelling, meaning that, in contrast with summarizing sequential patterns, the purpose is to predict future usage behaviour. Probabilities can be calculated by using zero-, first- and higher order markov chains (Scholz, 2016). These orders differ in that the next page-category is predicted based only on the current page-category that is visited by the user (first order), or on a combination of the current page-category in combination with the page-categories that the user was visiting before the current page-category (higher order).

Clustering.

In complex healthcare systems, it can be assumed that the users are heterogeneous, meaning that there is variation in their behaviour. It will almost never happen that all users have (almost) the same unique user sessions (Scholz, 2016). Therefore, these user sessions can be clustered based on the transitions between page-categories. With clustering, the user behaviour can be described in a global way (Butts, 2008), meaning that the focus is not on transitions between the distinct page-categories. Instead, it focuses on all transitions between page-categories, and subsequently, complete user sessions can be typified by their cluster name based on the similarity of these user sessions.

Aim of the study

The aforementioned three log file analyses are assessed using log files produced through a healthcare information system, in an attempt to answer the following question: ‘How may log file analyses contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?’. This question is then answered in a structured manner by formulating a sub-question for each aforementioned analysis. These sub-questions are (1) 'How may summaries of sequential patterns contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?', (2) 'How may heatmapping contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?' and (3) 'How may clustering contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?'.

Method

The research question was answered through two separate phases. In the first observational phase, results of summarizing sequential patterns, heatmapping and clustering were described, in order to explore how users' behaviour can be described on the transaction level. Subsequently, results of parts of the system developed by a single department were selected. These results were used in the second, qualitative phase of this study, by presenting these to the employees of this department. They only got to see the results of this selection, so that they would know and understand the content of these results. Through a semi-structured interview with developers, it was explored whether and how employees could use these results within the Agile development method, so as to increase the match between the healthcare system and its users.

Observational phase (1)

Before the observational phase of this study was carried out, permission was obtained from the Ethics Committee of the University of Twente regarding collecting log files and analysing these log files (see Appendix A). The log files were treated as confidential, and only kept on secured self-owned servers. The following sections show the method for data collection, data preparation and data analysis within this phase.

Data collection.

The log files were collected through Extendia. This is an Internet-based support system that serves as an extension of the general practitioner information system (GPIS). It provides general practitioners (GPs) with a comprehensive range of conceptually distinct services. It is a closed system, meaning that users are required to purchase a subscription before they can use it. When users log in directly via the browser, he or she will start within the ‘Declarations’ area. When a user logs in via another connected system (e.g. through signal sign on in GPIS), he or she doesn’t have to log in, and starts on a different page, depending on the link the user has clicked on.

Via the white menu at the top (see figure 2 and 3) users can download TeamViewer, open manuals for Extendia, and they can navigate to their own profile. Depending on which parts they purchased, users can use the services ‘Collaborate’, ‘Patients’, ‘Declarations’, ‘Dashboards’, and ‘Practice Web’. They can navigate to these services via the blue main menu at the top (see figure 2 and 3). Within these services, users can navigate to the related sub-functions. Below, we briefly describe these distinct services:

(1) Collaborate: Within collaborate users can consult another specialist to ask a question about a patient. Secondly, users can refer a patient to a mental health group, while keeping the opportunity to gain insight in the treatment of that patient.

(2) Patients: The patient-area is strongly connected to the collaborate area, and offers users the opportunity to gain insight in patients’ historical and current medication, and their measurement values. Moreover, historical care activities of the patient are clearly displayed in order of time on the patient-timeline.

(3) Declarations: The declarations-area includes four services for doing declarations, between different healthcare providers and insurers. Within this area, users do declarations, and these subsequently appear in the declaration overview. Next to this, the attention-page gives an overview of rejected declarations. On the page manual invoices users find and print (to send or provide) all patient invoices that are created in Extendia’s mute screen. Lastly, there is an option for creating passant invoices and third-party invoices.

(4) Dashboards: The dashboard area offers users the opportunity for GPs to benchmark their results. They can make reports for insurance companies, and reflect their results in comparison to colleagues on financial level, and on practice and patient data.

(5) Practice web: This part of the system offers functionalities to the users to support daily practice. The main parts are an address book, a to do list, and messages. In the address book contact details of the institutions and care providers who are often approached with regard to the care surrounding the patients or the practice can be saved and be seen by all employees. Secondly, with the to do list, actions can be assigned to the various employees and users can keep track of who, and when these tasks were carried out. Lastly, through the message component internal communication is digitally recorded.

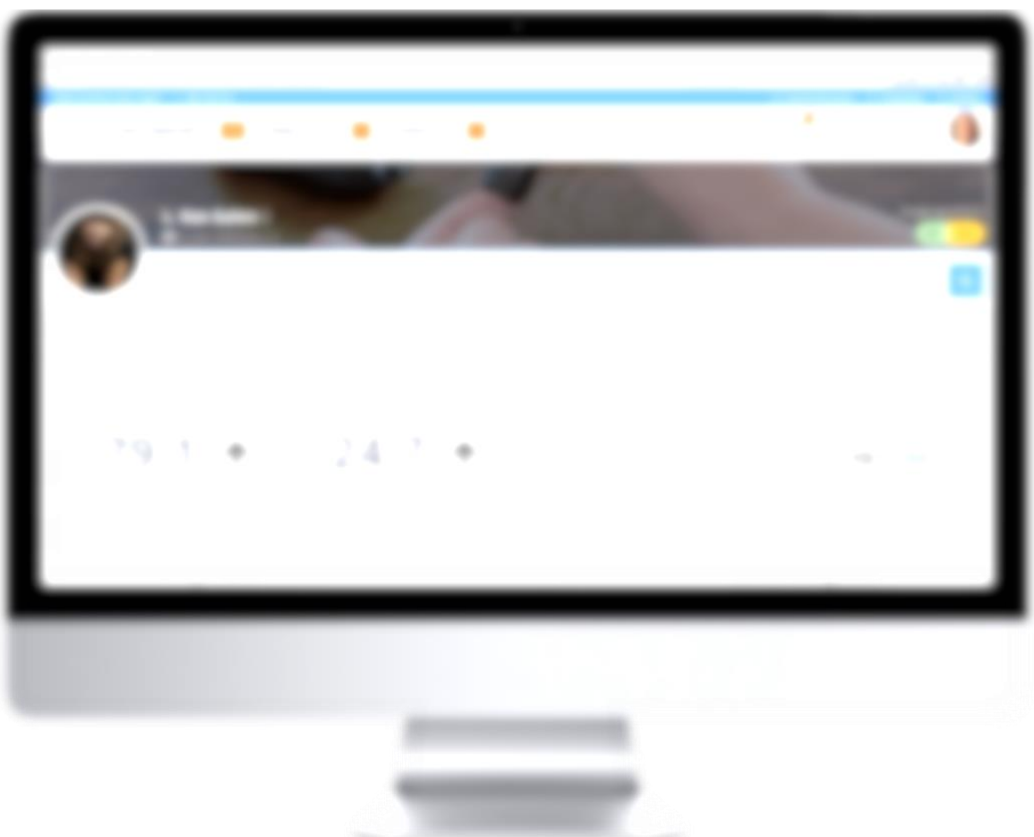

Figure 2. Screenshot of a fictitious page with measurement values of the Extendia system

Figure 3. Screenshot of a fictitious page with the timeline of a patient (Extendia)

Log files were collected through this system at server-side from the 18th of September, 2017 till the 17th of October, 2017. There was chosen for data of a one month period to fit well with the Agile development method (with a maximum duration of four weeks). To avoid drawbacks of logging data on the server side, the user name and a session-id were also logged by using a cookie. A user session ends when there is user inactivity for at least thirty minutes (Srivastava, Cooley, Deshpande, & Tan, 2000).

Data preparation.

There were 17924 user sessions obtained from Extendia that were collected from the 18th of September till the 17th of October, 2017. These were cleaned before user sessions were analysed. An example line of the log files collected at server-side can be seen below. This line of data consists of the date, the time, the method, the url requested, the status code, and the session identification number. In this specific example, the user requests the timeline within the patient-area for a specific patient on the third of September at around a quarter past four.

2017-09-03	16:19:59.5102	GET	/api/patient/xxxxxx/timeline	200	sessionID

Figure 4. One line of a log file collected at server-side

The columns date, time, url, and session id were used. To ensure privacy for the users, data were anonymized while keeping the opportunity to distinguish between sessions. Secondly, the url-variables were separated to allow for editing these variables separately in subsequent steps of cleaning the data. Background processes were removed, to ensure that log data only consisted of page-views initiated by the users. Moreover, when users visit a page for a specific patient, a unique page identifier is inserted. In order to analyse these pages for all users together, these identifiers were removed. At last, date and time were combined into one variable to use these data in detail.

As a subsequent step, each URL was pre-categorized into 62 page categories (i.e. "Upload Patient File", "Patient Timeline", "Address Book"), and grouped by their session numbers ordered by the variables date and time. This way, data of user sessions are formed. This is a set of views in a user session for a particular Web site (Cooley & Srivastava, 2000), between logging in and logging out to the system. The server session ends when there is user inactivity for at least thirty minutes. The cleaned data consists of comma-separated strings representing user sessions:

2017-09-18 11:24:05	SessionID	‘Insert Measurement Values for Patient’, ‘Patient Timeline’, ‘Patient Settings’, ‘Patient Treatment Plan’

Figure 5. User sessions as comma-separated strings

Data analysis.

In order to explore how users’ behaviour can be described on the transaction level, firstly, analyses related to the first research question were carried out. Therefore, sequential patterns in the log files were summarized. Descriptive statistics were calculated in Fluxicon Disco version 2.2.0 (Fluxicon, n.d.). Moreover, an overview of the usage was given in a process map.

Subsequently, the log files were analysed by use of RStudio version 1.0.143 (RStudio Team, 2016). In order to heatmap the log files as a first step for answering the second research question, firstly, a transition matrix was obtained which represents the probability of a transition between page categories. These transition probabilities were plotted in a heatmap.

Lastly, one randomly chosen day in the data was picked for the cluster analysis, due to limited computational power. These data were analysed in order to answer the third research question. The number of clusters was chosen based on the between and within sum of squares. These clusters were then typified based on the pages that mainly occur within user sessions that fall within the distinct clusters.

Qualitative phase (2)

After carrying out the observational phase of this study, these results were used during the second, qualitative phase of this study. Through a semi-structured interview, it was explored whether and how employees could use these results within the Agile development method, so as to increase the match between the healthcare system and its users.

Participants.

The participants were obtained through purposive sampling. An invitation was sent via a digital calendar to all employees of Developos who worked at the 'collaboration' department of Extendia. In total, ten employees worked at this department, of whom seven participated in this study. Their mean age was 31.14 (sd = 4.88) years, on a range between 25 and 39 years. Six participants were male, and one participant was female. Moreover, three participants were developers, two were designers, and two were analysts. Only Dutch speaking employees were included in this study.

Materials and procedure.

An interview protocol has been established and tested by means of a pilot project. In this pilot project, the interview was carried out with one employee (developer). A number of things have been adjusted on the basis of this pilot. Firstly, it was decided to also print the results, in addition to showing them via a beamer. This allows for the participants to better see what the results are. Secondly, it was decided to write down one striking and interesting result of every frequency overview and heatmap. This allows to keep the discussion about the results going when it falls silent.

After carrying out the pilot test, a meeting room was reserved within the company to carry out the focus group interview in a calm environment. The focus group interview took place on the 1st of November, 2017. At the beginning of the focus group interview, the participants were asked whether they agreed with recording the interview, and they were told that these data would be processed anonymously. All the participants agreed with this, and after that the recording of the focus group interview was started. The recording of the focus group interview was done with a mobile phone. In addition, notes were made during the interview which served as a backup in case the recording was not usable. Moreover, the most important comments were noted to look back on after the focus group interview.

Subsequently, the researcher introduced the focus group interview and results were shown. In the introduction, the participants were told that the results of the log file analyses were going to be presented, and that the focus was on increasing the match between the Extendia system and its users, and not on the performance of the system. A number of examples were given, and they were told from which period the data was collected. To increase their input, they were told that they could respond to the analyses that they were going to see, and there were no right or wrong responses. After this introduction, the results were presented on a beamer in the following order: (1) frequency overview ‘Collaborate’, (2) heatmap ‘Collaborate’, (3) frequency ‘Patient’ service, (4) heatmap ‘Patient’ service, and lastly, the (7) process map of all pages. These results are related to summarizing sequential patterns and heatmapping. It was decided to use these results, because these are related to the first and second research question. However, the clusters were not used during the interview, since this would make it necessary to include results related to parts developed by other departments. Participants were asked to explain how they interpret the results. At the end of the interview they were asked for which practical purpose they could use these results within the Agile process.

Analysis.

Prior to the analysis, the focus group interview was literally typed out using the F5 program. These transcriptions were imported into Atlas.ti version 8.1.3 (2017) to analyse the qualitative data. The data analysis took place by means of the coding method described by Boeije (2005): open coding and axial coding. The statements made by participants were evaluated by interpreting the meaning and assigning it a value code (open coding). A code has been assigned to relevant information per fragment. (Parts of) fragments received a maximum of one code. These codes were then grouped into overarching codes. Then, links were sought between the concepts, associations and combinations. This resulted in main groups and subgroups (axial coding).

The interview was double-coded: once by the researcher, and once by a fellow student of the researcher (10 percent of the data). This has been done to discuss the codes and to obtain a reliable coding scheme. Inconsistencies regarding the codes have been discussed, until a compromise was reached. For example, the code ‘Performance’ was removed based on this discussion, since this was not closely related to the research question. Lastly, the interrater reliability was calculated by means of Cohen’s Kappa in SPSS version 23.0.0.0.

Results

The results of the analysis performed are presented below, in two sub-sections. The first one focuses on the analyses – general results, summarizing sequential patterns, heatmapping, and clustering – performed within the observational phase of this study. The second sub-section focuses on the results from the qualitative part of this study, by describing the main categories and sub categories that appeared in the data obtained through the semi-structured interview.

Observational phase (1)

To illustrate analyses that were shown to the Agile team members, we describe the results of the observational phase of this study. In this phase of the study, log files that were collected server side were analysed. Firstly, a general overview of these log files was obtained and shown in the first paragraph. Subsequently, the sequential patterns were summarized, and described in the second paragraph. Thirdly, a heatmap of these log files was plotted, and lastly, user sessions were clustered to identify similar navigation patterns.

General results.

In total, 176678 activities were conducted within 17924 user sessions, an average of 9.86 activities in user sessions. In Figure 4, an overview of the distribution of these user sessions over the days is given. On weekdays, the number of user sessions is higher (m = 776.14) than during weekend days (m = 96.38).

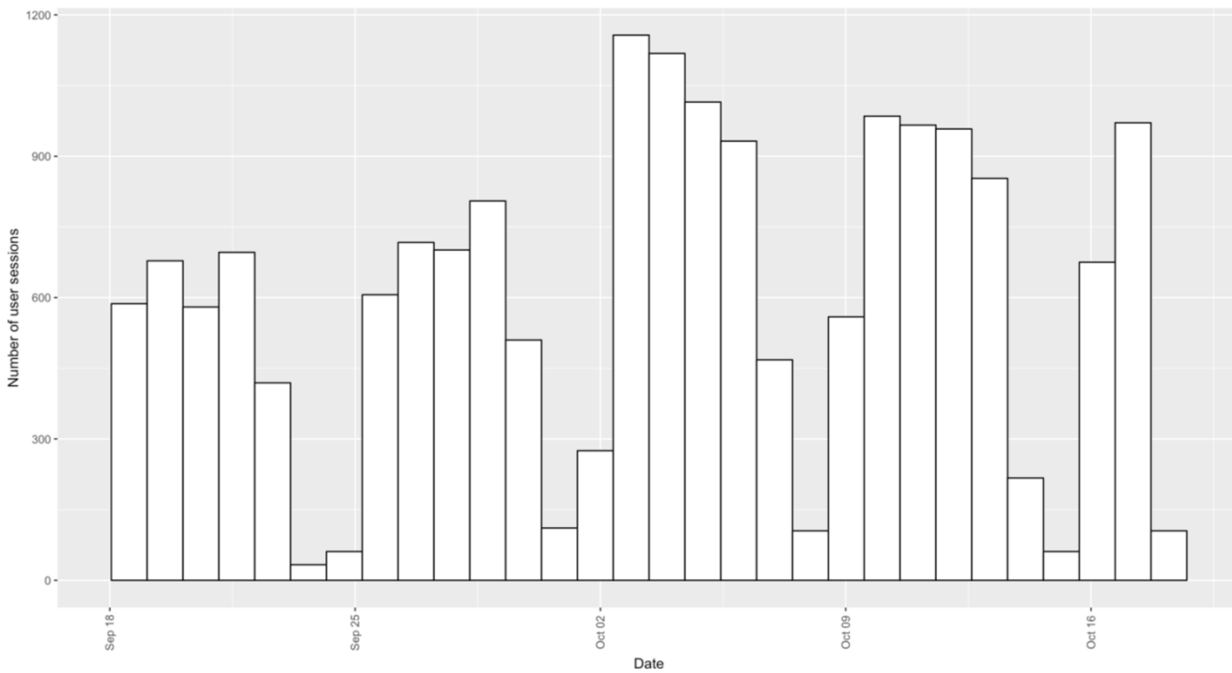

Figure 6. Number of user sessions over the days between September 18 and October 17, 2017

The page 'Patient timeline' was mainly visited within these user sessions (n = 23707), on which historical care activities of the patient are displayed in order of time. In addition, the page 'Insert measurement values' was then the most visited page (n = 21215). Table 1 shows the frequency and percentage of visits of the six pages that were visited most.

Table 1. *Most visited pages within Extendia system between September 18 and October 17, 2017. In the left column the page is shown, and the right column shows the number of times the page is visited with the percentage over the total number of activities.*

Page	**N times visited (%)**
Patient Timeline	23707 (13.42)
Insert Measurement Values for Patient	21215 (12.01)
Patient Settings	20306 (11.49)
Load Messages Collaborate	17284 (9.78)
View Report	15472 (8.76)
Patient Treatment Plan	14814 (8.38)

The least visited page was ‘Change conversation topic’ (n = 3). This page within ‘Collaborate’ is used for conversations between several healthcare providers. Here, the various care providers can consult each other about the care they give to a patient. Changing the subject of these conversations was done least within the Extendia system. Table 2 shows the frequency and percentage of visits of the five pages that were least visited.

Table 2. *Least visited pages within Extendia system between September 18 and October 17, 2017. In the left column the page is shown, and the right column shows the number of times the page is visited with the percentage over the total number of activities.*

Page	**N times visited (%)**
Add Participants	20 (.01)
Insight No. References to Mental Health Group	17 (.01)
To Do List	10 (.01)
Delete Participants	7 (.00)
Change Conversation Topic	3 (.00)

Summarizing sequential patterns.

In total, 6111 unique user sessions were found within all 17924 user sessions, meaning that on average 2.93 users have one unique user session. Overall, there is a lot of variation in how users navigate through the system. This section focuses on the paths that generally occur within user sessions. To begin with, the main unique user session occurred in 5.99 percent of all user sessions, and consisted of ‘Insert Measurement Values for Patient’, ‘Patient Timeline’ followed by the page ‘Patient Settings’ and lastly ‘Patient Treatment Plan’. To gain an understanding of the sequence of all user sessions, a process map was obtained in Fluxicon Disco (Fluxicon, n.d.). This process map can be seen in Figure 7, which shows 8.8 percent of all page-categories within the Extendia system, and 3.7 percent of the transitions between these pages.

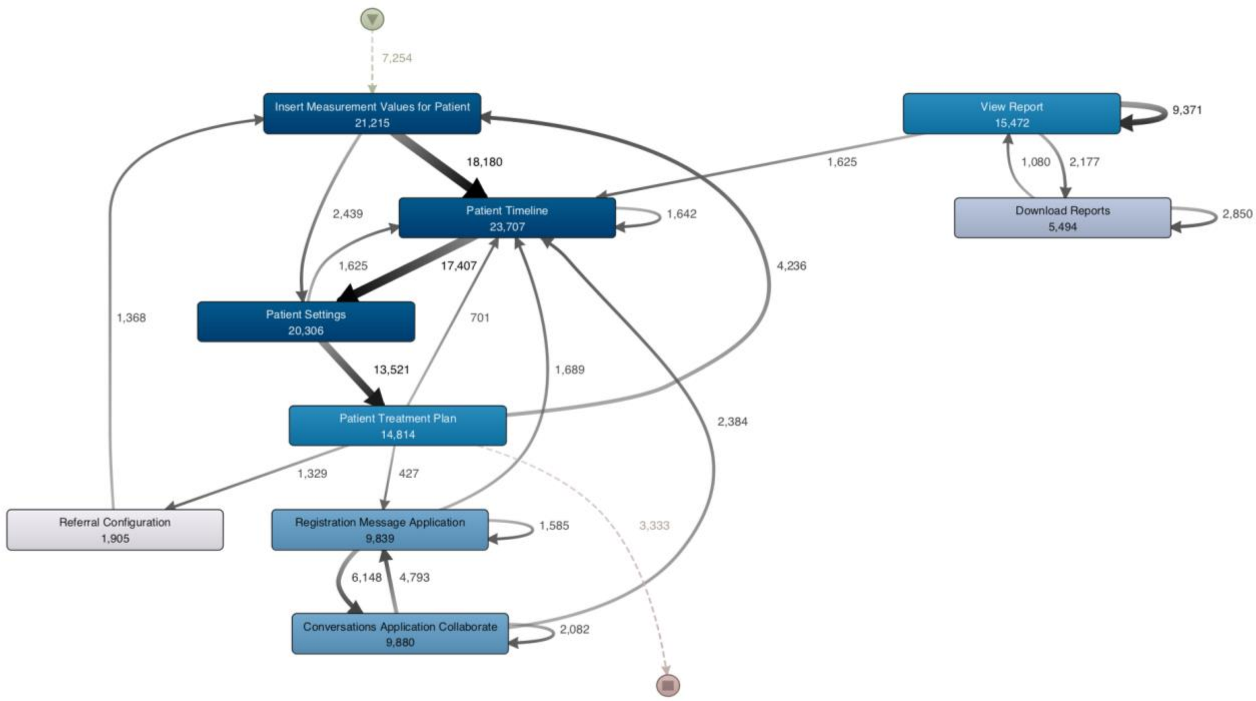

Figure 7. Overview of how users generally navigate through the Extendia system. The start of the process is illustrated by the triangle symbol at the top of the process map. Similarly, the end of the process is illustrated by the stop symbol. Pages are represented by boxes and the process flow between two pages is visualized by an arrow. Dashed arrows point to page-visits that occurred at the very beginning or at the very end of the process. The absolute frequencies are displayed in the numbers at the arcs and in the boxes. The thickness of the arrows and the colouring of the activities visually support these numbers.

In Figure 7, we can see that there are 7254 user sessions in the data set that all start with the page ‘Insert Measurement Values for Patient’. Moreover, users navigate to the page ‘Insert Measurement Values for Patient’ from the page ‘Patient Treatment Plan’ (n = 4236) and ‘Referral Configuration’ (n = 1368). Afterwards, the user sessions split into two alternative paths: In 18180 user sessions the page ‘Patient Timeline’ was visited after ‘Patient Treatment Plan’. The other

2439 user sessions navigate to 'Patient Settings' instead. In total, the page 'Insert Measurement Values' was visited most often (in total 21215 times) – which is more than the number of user sessions within the data (n = 17924). This comes from the dominant loop that goes indirectly through the 'Patient Timeline'. Repeatedly, the 'Patient Treatment Plan' is re-analysed while observing and inserting measurement values for the patient.

Another remarkable result in this figure is how and where the pages 'View Report' and 'Download Reports' occur in the user sessions. The page 'View Report' was visited 15472 times, which allows users to make reports for insurance companies, and reflect their results in comparison to colleagues on financial level, and on practice and patient data. From this page, 2177 users continue their path with visiting 'Download Reports'. This page ('Download Reports') was visited 5494 times, from which 1080 users visit 'View Report'. Similarly, after downloading reports within the Extendia system, 2850 users continue their path with downloading reports. There are a number of such patterns in the process map. For the readability of this study, these are not all mentioned in the text. The reader is referred to Figure 7 to see these patterns. Figure 7 also shows that, generally, user sessions end with visiting the page 'Patient Treatment Plan'.

Heatmapping log files.

As a first step for answering the second research question, ''How may heatmapping contribute to increasing the match between the healthcare system and its users within the Agile development method according to Agile team members?', the log files were plotted in a heatmap. Therefore, the log files were modelled with a first-order Markov chain. A heatmap was plotted for this transition matrix, with the y-axis representing the current page category and the x-axis the next page category. In Figure 8 the heatmap is shown, and in appendix B the second part of this heatmap can be found.

In Figure 8, we see that users often navigate to the pages ‘Insert Measurement Values’ and ‘Load Messages Collaborate’. Moreover, we see that there is a high probability that users repeatedly navigate from page ‘Address Book’ towards the same page. This page was visited for 133 times in a one month period. A page visited this small number of times relatively, would not appear quickly in a process map that gives an overview of the main visited pages.

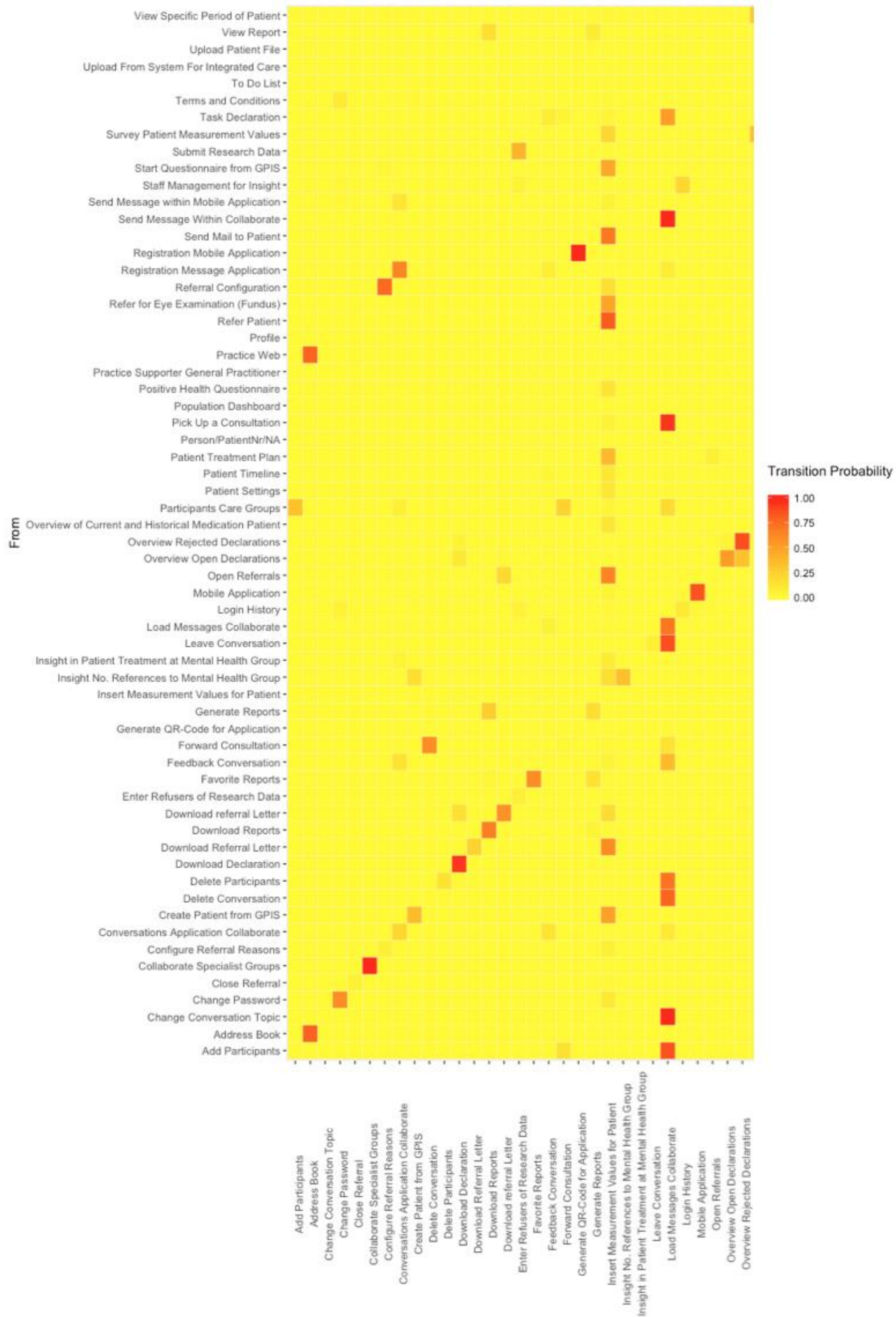

Figure 8. Heatmap of log files. The pages on the vertical line show from which page the user navigates and the horizontal line shows the pages toward which page a user navigates, where yellow squares stand for low transition probabilities, red squares for high transition probabilities, and orange squares for intermediate transition probabilities.

Clustering user sessions.

Lastly, the user sessions were divided into clusters with similar browsing patterns, as a first step for answering the third research question of this study. Firstly, one randomly chosen day in the log files was picked for the cluster analysis, due to limited computational power. These data were then divided into five clusters, based on the ‘elbow test’ (see appendix C). The clusters were typified based on the pages that appear in these clusters (see Table 3). The five clusters are named as the Information-seeking cluster (SS = 96.16), the Collaborative cluster (SS = 99.27), the Mixed cluster (SS = 193.40), the Administrative cluster (177.57), and the Patient-Oriented cluster (SS = 378.02). The total sum of squares within groups was 944.42 and the between sum of squares was 561.49. Appendix E shows the complete user sessions divided into five clusters.

Table 3. *Clusters with similar browsing patterns, along with the number of user sessions that were grouped into these clusters, and the three most visited pages within the clusters.*

Cluster	N user sessions within cluster	Most visited page	Second most visited page	Third most visited page
1. Collaborative	54	Load Messages Collaborate	Registration Message Application	Conversations Application Collaborate
2. Mixed	311	View Report	Registration Message Application	Conversations Application Collaborate
3. Administrative	15	Download Declaration	Overview Rejected Declarations	Download Referral Letter

Cluster	N user sessions within cluster	Most visited page	Second most visited page	Third most visited page
4. Patient-oriented	277	Patient Timeline	Patient Settings	Insert Measurement Values for Patient
5. Information-seeking	64	Download Reports	View Report	Population Dashboard

The first cluster, Cluster 1 (n = 54), shows a dominance of sessions involving collaboration with other care providers. The pattern of loading messages within ‘Collaborate’, followed by doing registrations for the message application and visiting this message application, shapes a picture of users looking at their notifications from other care providers and being actively involved in contacting these other providers.

Cluster 2, the Mixed Cluster (n = 311), shows a clear variance of events from event categories View Report, Registration Message Application, and Conversations Application Collaborate. As most user sessions belong to this cluster those can be interpreted to be the most typical sessions. This shows that the most typical user session involves a variety of tasks that can be performed within the system, meaning that users value to the wide range of tasks offered by the system.

Thirdly, Cluster 3 (n = 15), is the pattern that is least dominant in the user sessions. This cluster is typified by the administrative purposes that are performed within the sessions. The page-categories are related to downloading declarations, looking at an overview of rejected declarations, and downloading referral letters.

Cluster 4, the Patient-Oriented Cluster (n = 277), is also a very dominant cluster since it is the second most common among the user sessions. User sessions that fall within this cluster are oriented around the patient. Users are actively searching for a patient's timeline, change the settings of the patient, and enter measurements for that patient. This cluster seems to coincide with the layout of the system, of which a part is divided into searching and inserting patient-data.

The last cluster, Cluster 5, is the Information-Seeking Cluster (n = 64). This cluster shows a dominance of page-categories ‘Download reports’, ‘View Report’, and ‘Population Dashboard’. It is the most information behaviour related cluster as it reveals users searching for which patients have visited the GP, and what they have done there through reports.

Qualitative phase (2)

Four codes have been found in the qualitative data, namely ‘Hypotheses’, ‘Path comparison’, and ‘Benchmarking’. The number of times codes have been found in the data can be found in Table 3. The Cohen’s Kappa was calculated in SPSS to determine the interrater reliability of the main categories, and was found to be .81. In this section, these codes will be defined and the variance will be described using prime examples.

Table 3. *Main categories and sub categories found in the qualitative data.*

Main category (n)	Sub category	n
Hypotheses (34)	Usability problems	19
	User requirements	6
	Other systems	10
Path comparison (31)	Logical paths	13
	Illogical paths	12
	Incomplete paths	4
Benchmarking (22)	Page visits	19
	Users	3
Prioritize (17)	Redesign	4
	Making improvements	13

Hypotheses.

The code ‘Hypotheses’ was found 34 times in the data obtained through the semi-structured interview. Based on the results of the quantitative phase of this study, participants mainly communicated suspicions about the cause of these results. These beliefs of the participants were either related to usability problems, user requirements or the way of working in combination with other systems.

Usability problems were mentioned while discussing the heatmaps. Participants focussed mainly on the diagonal line of the heatmaps to think of usability problems. On this line participants could see how often a specific page is followed by loading the same page again. Examples of usability problems that were mentioned are lack of feedback, lack of clarity because the system is often adjusted, or entering incorrect passwords. An example of a usability problem related to the lack of feedback was mentioned by participant 1 (see Appendix B ‘Refer for Eye Examination’ to ‘Refer for Eye Examination’ for the results to which this quote relates to):

> *“Yes, we have just made a disable for that. Because Oculus received double or triple registrations. You had to press ‘sign up’ there and then you had to wait. But there was nothing to indicate that it was logging in at that time. That’s why they clicked again: they thought ‘nothing happens’.”* (R01, Q1:7)

Participants also formulated hypotheses which were related to user requirements, based on the heatmaps and frequencies of page-visits. Hypotheses about requirements are only related to functions that are unavailable within the system, and those related to existing functions belong to the ‘Prioritize’ code. User requirements related to non-existent functions are formulated based on the usage of existing functions. For example, respondents mentioned that deleting messages occurred much more often than leaving the conversation. Moreover, the heatmap revealed that after opening a conversation, users are inclined to directly delete the message. When a message is deleted, users cannot obtain the information about these messages anymore, while when leaving a conversation this possibility is retained. Thus, according to the participants, there is no need for an archive function. An example of a participant mentioning this is:

> *"Well, if you leave the conversation then you say: 'Ok, I have nothing to do with this conversation, but I want to be able to read it later'."* (R01, Q1:128)

Lastly, participants formulated hypotheses related to the way the system is used together with other systems, and reasons for this way of working with the other system. Usability problems are also formulated based on these hypotheses. Participants are focussed on the page from which users go to another system or the other way around. The movement towards another system indicates missing parts in the system. Moreover, when users move towards the system from another system, this is an indication for functions that users find valuable. This because they have to take extra steps from another system to use these functions. Moreover, based on the heatmap, participants form hypotheses based on the diagonal line. When there is a lot of repetition, users see another system as the starting point for carrying out their tasks. An example is the repetition of coming from a GPIS system. Based on this, the participants also formulate usability problems, e.g. there isn't a possibility to search based on maiden name. Participant 6 mentioned this (this quotation is related to a loop that occurred in the process map):

> *"The user is working from GPIS, and goes to Extendia via health portal with a button navigating to Extendia. When you do this a new tab opens with that patient information. And when you're done, you close that tab again and then you go to your GPIS again, where the user searches the next patient. From there the user goes back to the Extendia button and then Extendia opens again."* (R01, Q1:147)

Path comparison.

The code 'Path comparison' was found 31 times in the data obtained through the semi-structured interview. Based on the results of the quantitative phase of this study, participants compared the paths that occurred in the results of the quantitative phase of this study, with their expectations of how users navigate through the system. Paths are defined here as a sequence of requests to the server in time (as shown in the heatmap and process map). Paths that contain steps in other systems are excluded in this code, to avoid overlap with the code 'Hypotheses'. The comparison of paths in the results with their expectations were either logical, illogical or incomplete. Participants mentioned logical paths through the system, by which they mention that users carry out the steps in the right order to fulfil their tasks. This gives reason to state that certain parts of the system must be maintained in the same way. An example of a participant mentioning this is:

> *"But that is to be expected right. You start a conversation with someone. You get an answer, and then you're done."* (RO3, Q1:6)

Contrary to mentioning paths that match their expectations, participants also mentioned paths through the system that didn't match their expectations (process map and heatmap). Moreover, they mentioned results indicating that users carry out certain steps towards fulfilling a certain task, but the users don't carry out the last number of steps to fulfil this task. An example of such an incomplete path was mentioned by participant 3:

"But now it becomes very complicated because the user does not close the referral. At a given moment this patient still appears to be in the overview. Then that user calls us to ask how this referral can be closed, and then we have to search for a long time, and then we finally come to the conclusion that the user has not closed the referral at the end." (R03, Q1:41)

Benchmarking.

Thirdly, participants formulated thoughts related to the frequencies that were displayed regarding the use of the system. They wanted to compare these frequencies with other frequencies in order to obtain valuable clues about possible navigation problems within the system. These comments were coded as 'Benchmarking', and appeared 22 times in the qualitative data. Firstly, the participants compared the number of page-visits to another number of page-visits, to get an idea of the use relative to another function. For example, they compared the number of times the conversation topics were changed compared to the total number of messages that were opened. This gave indications for whether users find these functions valuable. Secondly, it seemed useful to the participants to compare frequencies of the use in the past with the current frequencies of use. For example, they could see whether recent developments concerning new users led to a logical increase in the number of page visits. An example of a participant mentioning this is:

> *"So, if you analyse the frequencies, there is an opportunity to see whether the usage increases enough when a new care group is added to the system. If not, this gives reason to think something is not going well with the implementation of the system for the new users"* (R01, Q1:24)

Prioritize.

Lastly, participants indicated opportunities for the results to be used for prioritizing within the Agile development process. This code, ‘Prioritizing’, occurred 17 times in the data. Firstly, they mentioned some cues in the results for prioritizing a redesign of the system. To illustrate, if users had to navigate to a frequently visited page via a drop-down menu, this would give priority to redesigning this page. Navigating to this page should then be possible via a main menu. This was also the opposite for less visited pages. Secondly, they mentioned that the frequency of page-visits provides clues for increasing or decreasing priority for making improvements to the system. An example of this is a participant mentioning:

> *“Well, I think this is a very interesting result, because my colleague is doing a lot to improve that page. That is adjusted every time, and uhhhh, that kind of things, but this page is almost never used.”* (R01, Q1:105)

Discussion

The goal of this study was to explore the usefulness of log file analyses within the Agile development method, so as to increase the match between the healthcare system and its users. For the investigation, log files were analysed as a sequence of events, and presented to Agile team members during a focus group interview, to make it look like natural situations that occur within the Agile process. We have found that analysing log files gives valuable insights into users' behaviour, allowing Agile team members to focus on the user of the system. On the basis of this, Agile team members can formulate suspicions that give rise to user-centred adaptations. However, the exploratory nature of this study calls for an investigation into the ways in which these user-centred adaptations can be integrated within Agile. After an exploration of our current findings, our discussion focuses on possible explanations for these findings, and the extent in which these findings can be applied in the steps of the Agile iterations. Finally, we discuss the limitations of our study, and subsequent recommendations are made, so that follow-up research can come to a more concrete description of the application of log file analysis within Agile.

Observational phase

The first phase of this research described the results of the log file analyses, in order to illustrate what kind of insights log file analyses provides. Here we relate the analyses to the insights obtained, based on the assessment of the researcher. Firstly, we illustrated the results of summarizing sequential patterns, which provided insight into the prominent order that occurs within user sessions. It seemed that the prominent order was related to the effectiveness with which users navigate through the system. For example, the loop between 'View Report' and 'Download Reports', indicated that users repeatedly download reports after having viewed them.

If the system would have offered an option to download several reports at once, these navigation paths could be more efficient.

Secondly, heatmapping log files provided insight into the transitions between lesser-visited pages within the system. Thus, this was a more detailed overview of the simplified representation of summarizing sequential patterns. The addition of the heat map is giving valuable prompts for possible bottlenecks in key functions within the system. It may be that a page is not visited often, but that it does have an important function. For example, the link between GPIS and Extendia only needs to be made once, but if this does not work, VIP live will be virtually unusable.

Thirdly, the clusters provided insight into groups of user sessions led by similar mental models of the system. The variation that exists between user sessions could be summarized in five clusters, creating an understanding of differences in user behaviour. In this way, the user behaviour can also be looked at in a more personalized way. We also saw a clear example of how mental models can change during the use of a system. For example, the 'Administrative cluster' was the least dominant cluster, while in the first instance the system only offered tasks related to administration.

Given these points, we have showed that describing users' behaviour on the transaction level can provide us valuable information about how users navigate through a complex healthcare system. This is in line with previous research related to eHealth technology (Han, 2011; Kelders, Bohlmeijer, & Van Gemert-Pijnen 2013; Sieverink et al. 2017).

Qualitative phase

Three concrete ways of representing log files were shown to Agile team members. To begin with, frequencies of page visits gave valuable prompts for prioritizing between possible adjustments, for stating hypotheses about user requirements, and to benchmark frequencies of use. We have seen that, to make the frequencies of use meaningful, Agile team members have to see it in relation to the use of a different part of the system or of a different point in time. By normalizing these frequencies, the results of the analyses can be adapted to this need for benchmarking. This can be done by for example using averages per user (Rodden, Hutchinson, & Fu, X., 2010).

Secondly, Agile team members were mainly interested in comparing expected paths with the actual paths in which users navigate through the system, while looking at summaries of sequential patterns. However, with log file data it is difficult to know what the user was trying to accomplish, urging the need for integrating the log files analyses with other methods for data collection about the user.

Thirdly, while looking at the heatmaps, Agile team members mainly focused on the diagonal line for stating hypotheses concerning usability problems. This might be because this part of the analysis provides the most valuable cues. On the other hand, a heatmap might be too excessive in information, forcing Agile team members to limit their focus on a single part of the analysis.

Lastly, the clusters could not be using during the semi-structured interview, because this user behaviour was described over the entire system. Within Agile, several Agile teams work on distinct parts of the software. Therefore, these sorts of analyses are difficult to use within an Agile process, which prevents employees from gaining a better understanding of user groups. By looking only at the previous analysis that is adapted to the parts they have worked on themselves, there is a risk that they will get a kind of tunnel vision, so that the complexity of user behaviour is overlooked.

In general, we have seen that participants focussed on specific parts of the analyses, indicating that the current analyses are too generic to track down specific cues related to UCD. A possible solution might be relating analyses explicitly to a goal, which is according to Rodden, Hutchinson and Fu (2010) a condition for a log file analysis to be useful. Within Agile, this means that before releasing added functionality, Agile team members decide the goals of this new software, point out which log data indicates success or failure of this goals, and translate this in specific metrics (Rodden, Hutchinon, & Fu, 2010). Parts of the current analyses are well suited for the purpose of serving as metrics. This way, the team members do not have to search for information from the analyses, but they can immediately see the relevant results after an iteration. This makes the use of log file analyses less time-consuming, making it easier to integrate within an Agile process with fast iterations. However, Agile team members can formulate valuable hypotheses based on the complete analysis. By linking it only to specific goals, the possibility to state hypotheses is reduced. Given that valuable hypotheses can be stated by analysing all user behaviour, we propose the use of goal-related analyses alongside the use of complete analyses within the Agile development process.

Study limitations

Having examined our findings, we now inspect in more detail the limitations of our exploratory study. Firstly, we discuss the subjective assessment of page-categories, and secondly the usefulness of clustering user sessions. For both limitations, it is discussed how this may have affected the results of the current study.

Subjective page-categories.

In this study, the urls from the log files were divided into page categories to keep the results clear. If urls were retained, results would become very extensive and unclear. This classification into page-categories was done on the basis of the subjective assessment of the researcher. Despite the benefits that this classification entailed, this categorization might have affected the sequences in which pages occurred in the log files. Ish-Shalom and Hansen (n.d.) suggested to use a similarity-based clustering technique for grouping these urls into categories as a solution to this problem. By using this technique for the division of urls into page-categories, the reliability of the study can be increased.

Usefulness of clustering user sessions.

In the qualitative phase of this study, the typified clusters were not shown to the participants during the interview. This was because only parts of the system that the users had worked on were included, to increase the participants' understanding of the results during the semi-structured interview. However, including this in an interview in which the practical applications of these results are explored, might give more information about the usefulness of this cluster analysis. Therewith, a first step towards validating users' mental models can be done for the cluster analysis. One step further towards validating these mental models can be done by involving users in the study.

Recommendations for future research

We will now discuss how the current study can be continued in the future. Two recommendations are made for follow-up research and a first brief outline of how such research can be addressed. Finally, we conclude with a broader description of the contribution of this work.

Level of detail.

In the current study, the level of detail in which insight was gained into the user behaviour was chosen arbitrarily. For example, more pages can be included in the process maps, as well as the links between these pages. In the current study, it was found that the analyses on the transaction level were useful within the Agile process, but a different level of detail in the analyses could provide more useful insights to increase the match between the system and the user. By using different levels of detail in a subsequent study, it can be determined which level of detail provides the most useful insights. To make it more concretely, the current analyses can be shown at various levels of detail to the various departments within the software company, where each team is asked what they can do with these results. Then these results can be compared to determine which results have yielded the most valuable insights. By doing this within a single company, the results are more comparable than when analyses of different systems are carried out at different levels of detail.

Research in EHealth interventions.

Currently, a growing number of research focuses on explaining the effects of EHealth interventions (e.g. Sieverink, 2017). According to Catwell and Sheikh (2009) evaluating EHealth interventions while they are being designed, developed and deployed is important to minimize the risks, and to maximise the potential benefits of EHealth interventions. The use of log files has the potential to continuously evaluate the use of EHealth interventions (Sieverink, 2017). Moreover, it has the potential to explain what has been called the ‘black-box’ of EHealth interventions (Sieverink, 2017). This term refers to the problem that the presence or absence of an eHealth intervention is used to explain the effects, while it is not clear how this intervention was precisely used (Sieverink et al., 2017). The methods used for analysing log files in the current study can be used in EHealth research to describe users’ behaviour within an EHealth environment. In future research, the extent to which an intervention has an effect can be correlated with a description of user behaviour as was done in the current study. In this way, the description of user behaviour is supplemented with sequential statistics, over the more classical use of descriptive statistics.

Due to the exploratory nature of this study, findings might not have attained a more concrete answer on the exact added value of log file analyses within the Agile process. We did however pave the way towards an approach to increase the focus on the user within the Agile process, that can be more easily integrated within Agile then the more classical approaches to UCD.

References

Agile. (n.d.). In *OxfordDictionaries.com*. Retrieved from https://en.oxforddictionaries.com/definition/agile

Ardito, C., Buono, P., Caivano, D., Costabile, M. F., & Lanzilotti, R. (2014). Investigating and promoting UX practice in industry: An experimental study. *International Journal of Human-Computer Studies*, *72*(6), 542-551. doi: 10.1016/j.ijhcs.2013.10.004

Ardito, C., Buono, P., Caivano, D., Costabile, M. F., Lanzilotti, R., Bruun, A., & Stage, J. (2011). Usability evaluation: A survey of software development organizations. In *Proceedings of the 23rd international Conference on Software Engineering and Knowledge Engineering (SEKE 2011)* (pp. 282-287). Knowledge Systems Institute Graduate School.

ATLAS.ti version 8.1.3 [Software]. (2017). Retrieved from https://atlasti.com/

Bak, J. O., Nguyen, K., Risgaard, P., & Stage, J. (2008). Obstacles to usability evaluation in practice: A survey of software development organizations. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (pp. 23-32). ACM

Boeije, H. (2005). *Analyseren in kwalitatief onderzoek* (2nd ed.). Den Haag, The Netherlands: Boom Lemma Uitgevers.

Brhel, M., Meth, H., Maedche, A., & Werder, K. (2015). Exploring principles of user-centered agile software development: A literature review. *Information and Software Technology*, *61*, 163-181. doi: 10.1016/j.infsof.2015.01.004

Butts, C.T. (2008). Social network analysis: A methodological introduction. *Asian Journal of Social Psychology, 11*, 13–41. doi: 10.1111/j.1467-839X.2007.00241.x

Catwell, L. & Sheikh, Z. (2009). Evaluating eHealth interventions: The need for continuous systemic evaluation. *PLOS Medicine, 6*(8), e1000126. doi: 10.1371/journal.pmed.1000126

Dumais, S., Jeffries, R., Russell, D. M., Tang, D., & Teevan, J. (2014). Understanding user behavior through log data and analysis. In *Ways of Knowing in HCI* (pp. 349-372). Springer: New York, NY.

Dybå, T. & Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and Software Technology, 50*(9), 833-859. doi: 10. 1016/j.infsof.2008.01.006

Fluxicon (n.d.). *Discover your processes.* Retrieved June 13, 2018 from https://fluxicon.com/disco/

Gemert-Pijnen, van, J.E., Nijland, N., Limburg, van, M., Ossebaard, H.C., Kelders, S.M., Eysenbach, G., & Seydel, E.R. (2011). A holistic framework to improve the uptake and impact of eHealth technologies. *Journal of Medical Internet Research, 13*(4), [e111]. doi: 10.2196/jmir.1672

Han, J.Y. (2011). Transaction logfile analysis in health communication research: Challenges and opportunities. *Patient Education and Counselling, 82*(3), 307-12. doi: 10.1016/j.pec.2010.12.018

Interaction Design Foundation (n.d.). User Centered Design. Retrieved from https://www.interaction-design.org/literature/topics/user-centered-design

Ish-Shalom, S. & Hansen, S. (n.d.). Visualizing clickstream data as discrete-time markov chains. Retrieved from https://web.stanford.edu/class/archive/cs/cs448b/cs448b.1166/cgi-bin/wiki/images/a/a0/Hansen_final_report.pdf

Jokela, T., & Abrahamsson, P. (2004). Usability assessment of an extreme programming project:

Close co-operation with the customer does not equal to good usability. In *International Conference on Product Focused Software Process Improvement* (pp. 393-407). Springer.

Kelders, S.M., Bohlmeijer, E.T., & Van Gemert-Pijnen, J.E. (2013). Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. *Journal of Medical Internet Research, 15*(8), [e172]. doi: 10.2196/jmir.2258

Kushniruk, A. (2002). Evaluation in the design of health information systems: Application of approaches emerging from usability engineering. *Computers in Biology and Medicine, 32*(3), 141-149. doi: 10.1016/S0010-4825(02)00011-2

Matharu, G. S., Mishra, A., Singh, H., & Upadhyay, P. (2015). Empirical study of agile software development methodologies: A comparative analysis. *ACM SIGSOFT Software Engineering Notes*, *40*(1), 1-6. doi: 10.1145/2693208.2693233

Moreno, A.M., Seffah, A., Capilla, R., & Sanchez-Segura, M.-I. HCI practices for building usable software. *Computer*, *46*(4), 100–102. doi: 10.1109/MC.2013.133

Neisser, U. (1976). Cognition and reality: Principles and implications of cognitive psychology. New York, NY: W. H. Freeman.

Newell, A. (1982). The knowledge level. *Artificial Intelligence, 18*, 87-127.

Newell, A. "Putting it all together." (Chapter 15) In Klahr, David, and Kenneth Kotovsky, eds. Complex information processing: The impact of Herbert A. Simon. Psychology Press, 1989

Nguyen, L., Bellucci, E., Nguyen, L.T. (2014). Electronic health records implementation: An evaluation of information system impact and contingency factors. *International Journal of Medical Informatics, 83,* 779–796. doi: 10.1016/j.ijmedinf.2014.06.011

Nielsen, J. (2010). Mental models. Retrieved September 30, 2017, from https://www.nngroup.com/articles/mental-models/

Norman, D.A. (2014). Some observations on mental models. In D. Gentner & A.L. Stevens (Eds.), *Mental Models* (pp. 15-22). New York, NY: Psychology Press.

Norman, D.A. (2013). *The design of everyday thing: revised and expanded edition.* New York, NY: Basic Books (AZ).

Ortega-Navas, del, M.C. (2017). The use of new technologies as a tool for the promotion of health education. *Procedia Social and Behavioral Sciences, 237,* 23-29. doi: 10.1016/j.sbspro.2017.02.006

Ratwani, R.M., Fairbanks, R.J., Hettinger, A.Z., & Benda, N.C. (2015). Electronic health record usability: Analysis of the user-centered design processes of eleven electronic health record vendors. *Journal of the American Medical Informatics Association*, *22(6)*, 1179–1182. doi: https://doi.org/10.1093/jamia/ocv050

Ries, E. (2011). *The Lean startup: How constant innovation creates radically successful businesses.* London, UK: Portfolio Penguin.

RStudio Team (2016). RStudio: Integrated Development for R (Version 1.0.143) [Computer software]. Boston, MA. Retrieved October 5, 2017. Available from http://www.rstudio.org/

Rodden, K., Hutchinson, H., & Fu, X. (2010). Measuring the user experience on a large scale. Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI ’10. doi:10.1145/1753326.1753687

Rogers, Y., Sharp, H., & Preece, J. (2015). *Interaction design: Beyond human-computer interaction*. Chichester, UK: John Wiley & Sons.

Russell, D. M. (2014). Understanding user behavior at three scales: The AGoogleADay story. Retrieved from http://pub-tools-public-publication-data.storage.googleapis.com/pdf/44006.pdf

Scholz, M. (2016). R package Clickstream: Analyzing Clickstream data with markov chains. *Journal of Statistical Software, 74*(4), 1-17. doi: 10.18637/jss.v074.i04

Schwaber, C. (2006). The root of the problem: Poor requirements. *Forrester Research*, 1–7.

Schraagen, J. M. (2017). Naturalistic Decision Making. In L. J. Ball & V. A. Thomson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 487-501). New York, NY: Routledge.

Sieverink, F. (2017). *Opening the black box of EHealth: A mixed methods approach for the evaluation of personal health records* (dissertation). Retrieved from https://ris.utwente.nl/ws/portalfiles/portal/19226592

Sieverink, F., Kelders, S., Poel, M., & Gemert-Pijnen, van, L. (2017). Opening the black box of electronic health: Collecting, analyzing, and interpreting log data. *JMIR Research Protocols, 6*(8), [e156]. doi: 10.2196/resprot.6452

Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *SIGKDD Explorations, 1*(2), 12-23. doi: 10.1145/846183.846188

Sohaib, O. & Khan, K. (2010, July). Integrating usability engineering and agile software development: A literature review. Paper presented at the *International Conference on Computer Design and Applications*, Qinhuangdao. Retrieved from https://pdfs.semanticscholar.org/1b8c/9ff550845d61a60fcdeffa71417fde5ef26c.pdf. doi: 10.1109/ICCDA.2010.5540916

Tang, P.C., Ash, J.S., Bates, D.W., Overhage, J.M., & Sands, D.Z. (2006). Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association, 13*(2), 121-126. doi: 10.1197/jamia.M2025

The Standish Group. (2014). The Standish group: The chaos report. *Project Smart*, 1-16. doi: 10.1016/S0895-7061(01)01532-1

Appendix A. Permission ethics committee

UNIVERSITY OF TWENTE.

Faculty of Behavioural Sciences

ETHICS COMMITTEE (CE) FACULTY OF BEHAVIOURAL SCIENCES.

APPLICATION FORM INTENDED RESEARCH PROPOSALS ETHICS COMMITTEE BEHAVIOURAL SCIENCE, VERSIE 2

1. Background human research participants

1. Does your research proposal concern medical-scientific research?

NB: Medical-scientific research in this context is defined as 'research which is carried out with the aim of finding answers to a question in the field of disease and health (etiology, pathogenesis, signs/symptoms, diagnosis, prevention, outcome or treatment of disease), by systematically collecting and analysing data. The research is carried out with the intention of contributing to medical knowledge which can also be applied to populations outside of the direct research population.'

No

2. Title

2b. Date of the application

19-09-2017

2a. What is the title of the research (max. 50 characters)?

PLEASE BE AWARE: If you are going to make use of the SONA system, at the 2nd question you have to fill in the same title that you will be using in SONA. This title will (in case SONA is used) also be visible to the human research participants..

Log File Analysis in the Health Informatics Domain

3. Contact information Reseachers / Conductors of the research

3a. Initials

I.

3b. Surname

ten Klooster

3c. Research Department (if applicable)

0

3d. Student number

1795562

3e. Email address

i.tenklooster@student.utwente.nl

3f. Telephone number (during research):

0620730557

3g. If there is more than one conductor of the research, than please fill in the text boxes below and fill in the contact information (initials/surname/email address/phone number) of all the conductors of the research.

-

Appendix B. Heatmap of log files (2)

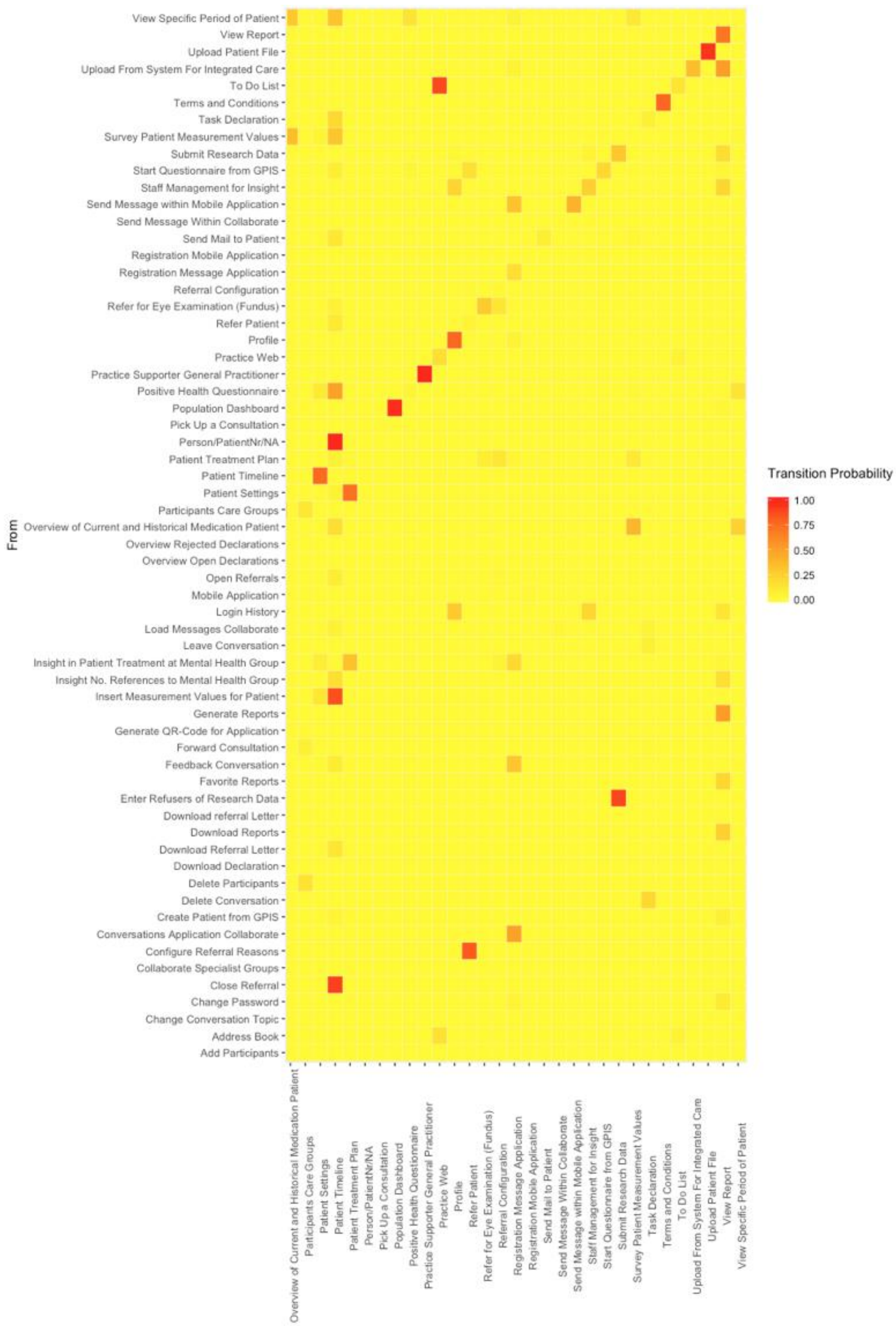

Figure 9. Heatmap of log files (2)

Appendix C. Table with sum of squares per *n* clusters

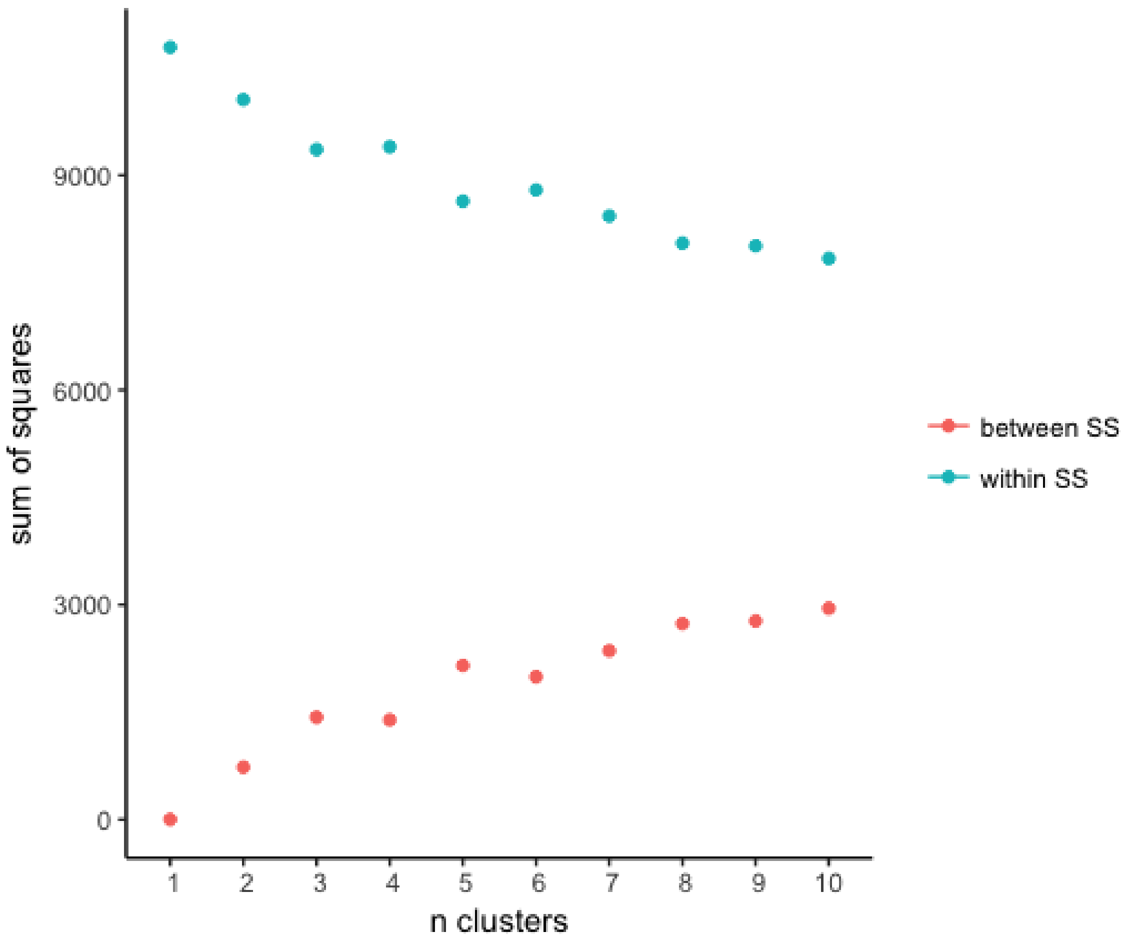

Figure 10. Table with sum of squares per *n* clusters

Appendix D. R Code

```
------------------------------------------------------------------------------------------

  # Load packages

------------------------------------------------------------------------------------------

library(readxl)

library(data.table)

library(clickstream)

library(arulesSequences)

library(ggplot2)

library(plyr)

library(tidyverse)


------------------------------------------------------------------------------------------

  # General results

------------------------------------------------------------------------------------------

# Graph with number of sessions per day

unique_session <- subset(Remove5, !duplicated(Remove5$SessionId))


ggplot(data = unique_session, aes(x = unique_session$DateTime)) +

  geom_histogram(color="black", fill="white") +

  xlab("Date") +

  ylab("Number of user sessions") +

  theme(axis.text.x = element_text(angle = 90, vjust = 0.5)) +
```

```
  theme_bw()


# Make a Clickstream object
Analysis_UserSessions <-
  as.list(aggregate(Remove5$AreaControllerFunction ~ Remove5$SessionId, Remove5,
function(x) c(as.character(x))))


cls <- Analysis_UserSessions$`Remove5$AreaControllerFunction`
class(cls) <- "Clickstreams"


# Summarizing user sessions
summary(cls)


-----------------------------------------------------------------------------------------
  # Heatmapping log files
-----------------------------------------------------------------------------------------
mc <- fitMarkovChain(clickstreamList = cls, order = 1, control = list(optimizer = "quadratic"))
mc
summary(mc)
plot(mc, order = 1)


hmPlot(mc, order = 1, absorptionProbability = FALSE, title = "Heatmap transitions",
       lowColor = "yellow", highColor = "red", flip = FALSE)
```

```
-----------------------------------------------------------------------------------------

  # Clustering log files

-----------------------------------------------------------------------------------------

maxOrder <- 5

result <- data.frame()

for (k in 1:maxOrder) {

  mc <- fitMarkovChain(clickstreamList = cls, order = k)

  result <- rbind(result, c(k, summary(mc)$aic, summary(mc)$bic))

}

names(result) <- c("Order", "AIC", "BIC")

result


clusters <- clusterClickstreams(clickstreamList = cls, order = 1, centers = 5)

clusters

summary(clusters)

cluster1 <- clusters$clusters[[1]]

cluster2 <- clusters$clusters[[2]]

cluster3 <- clusters$clusters[[3]]

cluster4 <- clusters$clusters[[4]]

cluster5 <- clusters$clusters[[5]]
```

Appendix D. User sessions divided into clusters

[[1]] Collaborative cluster

001: Patient Timeline Insert Measurement Values for Patient Patient Settings Patient Treatment Plan View Report View Report View Report Download Reports Insert Measurement Values for Patient Patient Timeline Patient Settings Patient Treatment Plan/

[excluded]