TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies

Kristian Talviste 192583YVEM

Using Process Mining to Evaluate the Data Capturing Optimization on the Example of Psychiatric Decision Support Tool DocuMental

Master's thesis

Supervisor: Janek Metsallik
MSc, lecturer

Tallinn 2023

TALLINNA TEHNIKAÜLIKOOL
Infotehnoloogia teaduskond

Kristian Talviste 192583YVEM

Protsessikaevanduse abil andmesisestuse optimeerimise võimaluste hindamine psühhiaatrilise otsustustoe süsteemi DocuMental näitel

Magistritöö

Juhendaja: Janek Metsallik
M.Sc, õppejõud

Tallinn 2023

Author’s declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Kristian Talviste

11.05.2023

Abstract

Background: Diagnoses for mental and behavioural disorders are based on data gathered via structured clinical interviews, consisting of extensive questions and observations about the patient. In psychiatry, this has traditionally been done using pen-and-paper methods. Digital platforms have been introduced as a solution to this problem. However, there needs to be more evidence of the usefulness of these platforms. DocuMental is a modern clinical decision support system currently being developed in Estonia, which is designed to aid in the diagnosis, treatment, and management of mental disorders and provides a platform for digitized clinical workflows in mental health. **Aim**: This research aims to evaluate options for improving the data collection process in psychiatric decision support on the example of DocuMental. **Methods**: The research uses process mining techniques to assess the data structures and event logs in DocuMental. Various exclusion criteria are used to improve the data quality. **Results**: The research parsed and analysed ten files. Within the data structure files, 364 diagnoses, 889 questions, 3414 answer options between questions, and 21 749 interconnections between these structures were discovered. Additionally, the research identified 814 encounters. 58 978 actions were observed within the event logs of DocuMental for these encounters. The research analysed general statistics for 151 encounters due to data exclusion criteria and 39 encounters in depth within the anxiety checklist. The research analysed the psychiatrist's actual user journey and the theoretical shortest paths to reach anxiety-related diagnoses to observe if an excess number of questions were answered. In addition, 188 unused questions and 9232 actions of data duplication were identified. Furthermore, screen transitions and the time spent after diagnosis confirmation were analysed. **Conclusions**: The analysis done during this research indicated a 70,27% time-saving potential in the data collection process of DocuMental. Various metrics analysed show additional underperformance in reaching diagnoses as numerous unused questions and data duplication exists within the application.

This thesis is written in English and is 64 pages long, including 5 chapters, 5 figures and 7 tables.

Annotatsioon

Taust: Vaimse tervise ja käitumishäirete diagnoosimine põhineb struktureeritud kliiniliste intervjuude käigus kogutud andmetel, mis koosnevad suuresti küsimustest ja vaatlustest patsiendi kohta. Psühhiaatrias on seda traditsiooniliselt tehtud paber-ja-pliiats meetoditega. Infotehnoloogiliste arengutega on digitaalsed platvormid muutunud võimalikeks lahendusteks andmekogumisprotsesside moderniseerimiseks, kuid puuduvad tõendusmaterjalid nende platvormide kasulikkuse kohta. DocuMental on kaasaegne kliinilise otsustustoe rakendus, mis on hetkel arendamisel Eestis. Selle eesmärgiks on aidata vaimsete häirete diagnoosimisel, ravimisel ja ravi juhtimisel ning pakkuda platvormi vaimse tervise digitaliseeritud kliiniliste töövoogude jaoks. **Eesmärk**: Lõputöö eesmärgiks on hinnata võimalusi andmekogumisprotsessi parendamiseks psühhiaatrilises otsustustoe rakenduses DocuMentali näitel. **Metoodika**: Lõputöö kasutab protsesside kaevandamise metoodikat rakenduse DocuMental andmestruktuuride ja sündmuste logide hindamiseks. Andmekvaliteedi parandamiseks kasutatakse erinevaid välistuskriteeriume. **Tulemused**: Lõputöö käigus analüüsiti 10 faili. Andmestruktuure sisaldavate failide seast avastati 364 diagnoosi, 889 küsimust, 3414 vastusevarianti ning 21 749 eelnimetatud struktuuride vahelist seost. Lisaks tuvastas lõputöö 814 visiiti. DocuMentali sündmuste logidest tuvastati nende visiitide jooksul 58 978 toimingut. Võttes arvesse andmete välistamise kriteeriume, analüüsiti üldist tegevuste statistikat 151 visiidi puhul ning süvitsi analüüsiti 39 ärevushäirete diagnoosiga visiiti. Lõputöö analüüsis psühhiaatri tegelikku kasutaja teekonda rakenduses ja teoreetilisi lühimaid teekondi ärevushäiretega seotud diagnoosideni jõudmiseks, et tuvastada küsimuste liigvastamist. Täiendavalt tuvastati 188 kasutamata küsimust ja 9232 vastuste duplitseerimise toimingut. Lisaks analüüsiti ekraanisiirdeid ja aega, mida kulutati rakenduses peale diagnoosi kinnitamist. **Järeldused**: Selle lõputöö käigus tehtud analüüs näitas, et DocuMentali andmete kogumisprotsessis on 70,27% aja kokkuhoiu potentsiaali. Erinevad analüüsitud mõõdikud näitavad puuduseid diagnoosideni jõudmisel, kuna rakenduses esineb kasutamata küsimusi ja andmete korduvsisestust.

Lõputöö on kirjutatud inglise keeles ning sisaldab teksti 64 leheküljel, 5 peatükki, 5 joonist ja 7 tabelit.

List of abbreviations and terms

AI	Artificial Intelligence
DSM or DSM-5	The Diagnostic and Statistical Manual of Mental Disorders handbook
ICD or ICD-10	International Classification of Diseases – 10th Revision
eMH	e-Mental Health
MHapps	Mental Health applications
CDA	Computerized Decision Aid
EMR	Electronic Medical Records
FHIR or HL7-FHIR	Fast Health Interoperability Resources
DSS	Decision Support System
CDSS	Computerized Decision Support System

Table of contents

List of figures

List of tables

1. Introduction

We live in a data-driven world. Every person is estimated to generate around 1.7 megabytes of data each second in 2020 [1]. This amounts to about 146 gigabytes of data per day for every person. With almost 4.7 billion active internet users worldwide in October 2020 [2] it is estimated that around 2 500 000 terabytes of data are generated daily [3]. This data is used for different decisions in the world, some in real-time. The amount of data that has been amassed is nearly impossible for a human to go through manually. Luckily, technology has advanced rapidly in the last decades, and now many say that we have one of the most powerful tools ever available – artificial intelligence [4]. Big data would be useless without AI, and AI algorithms work better with a broad set of underlying data. The two complement each other [5]. All this data is gathered from different sensors or various devices and applications and is mainly formatted as machine-readable.

Psychiatrists gather subjective data daily [6]. In psychology, the patients tell their psychiatrists about their concerns, what they think about, what dreams they see, and if they are in pain. Things like that cannot be measured, so they cannot precisely be written down to a machine-readable structure for an AI to use. It is also impossible to map these findings to exact biomarkers due to the complexity of these biological findings and the very nature of mental and behavioural disorders. A study by García-Gutiérrez tried to link biomarkers with certain neuropsychiatric diseases. However, none were accurate or useful in practice [7]. Even though technology has been used in psychology for decades [8], and is being commonly used by psychiatrists in practice, it can be too early for a conclusion regarding the usage of technology on the quality of care [9]. A Forbes article on Kintsugi, an AI application designed to detect depression in a person’s voice, is said to be 80% clinically accurate when determining mental health issues compared to 47,3% accuracy by professionals [10], which is a significant improvement and gives a solid case to use the support of an AI in the psychiatric decision process.

Psychiatry is still in baby steps when using technology, mainly due to the assessments being more intuition-based and heuristic-based decisions [11]. Diagnosing mental illnesses uses the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) handbook, published in 1952 and revised later. Section III of the DSM-5 lists assessments to go through during an interview with patients to diagnose mental disorders [12]. Another guideline comes from the ICD-10 Classification of Mental and Behavioural Disorders, founded around the 1960s by the World Health Organisation. This guideline includes a range of mental conditions and their descriptions. The descriptions state conditions to be present and absent for a diagnosis, including markers and diagnostic guidelines [13].

However, with the growing popularity of patient applications, there are few advancements on the professional side. To the best of the authors' knowledge, not much is known regarding the tools psychiatrists use apart from the usual national or local applications used for electronic filing to fill out EMR-s. In one case, it is reported that many of these applications go against the guidelines and principles available to healthcare professionals [14]. Since many psychiatrists use software not specialized for their profession, or sometimes even the traditional pen-and-paper method, filling out necessary forms can be time-consuming. A research that looks at mental health apps in psychiatric treatment from the patient's perspective includes a complaint from a participant. The patient, who has had bipolar disorder since childhood, expressed that she sees potential in MHapps. However, they complement the core of good mental health care, which she considers to be personal attention from someone who cares. She expressed that personal encounters with medical professionals must stay more prominent than using any application. In her words, many medical visits have now become very computer or data focused, rather than looking at the patient's suffering in the room. Patients, such as herself, with mental disorders need this focus, empathy, and commitment to help solve their problems or alleviate their suffering, not spend their in-person sessions with the professional behind a computer screen or fiddling with devices [15].

1.1.Shortfalls of medical data collection processes

However, data in healthcare seems to be problematic. Many healthcare professionals are not technical people and need help with all the technological advancements that are going on, starting with new devices and software and ending with the various interfaces used. A professor of medicine at the University of California, Robert M. Wachter has published a book called “The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine’s Computer Age.” He takes a look at what problems technology has caused in healthcare. He calls healthcare the most information-intensive industry, which trails behind the rest of the industries with technological advancements and is having problems due to inconsistent data. Wachter describes the research and work of Matthew Burton at the Mayo Clinic, where they tracked how nurse practitioners gather data. What they found was astonishing. The nurses had to log in to 11 different information systems to gather what they needed. The data collection process took over 600 clicks and over 200 screen transitions. Whenever a new page comes up, the professionals must process the new visual information with their brains, adding extra time to the whole process. Burton went forward and, with his team, developed a screen that could be filled in with just 25 clicks. The time required for collecting the data during multidisciplinary rounds was reduced from 35 minutes to less than 5 [16]. This freed up time to deal with the patients, and because of this, fewer mistakes were made by the practitioners. Wachter also brings out that computers cause reduced attention to patients during in-person meetings, during which physicians spend most of their time working on computers, not on the patients [16]. All computer systems can have glitches, and healthcare is no exception. However, glitches in the computer systems in healthcare could kill patients. Due to possible errors, some physicians have called to abandon technology in healthcare and return to pen-and-paper-based processes. However, Wachter brings out that despite the problems, the evidence shows that healthcare is better and safer with computers than without them [16].

Another research done by Anni Männil highlights that, on average, 36.8% (5.33 minutes) of a visit time to a family physician is spent on computer-based tasks. These include searching for patient electronic health/medical records, authenticating to different systems and services, clinical documentation, writing referrals, issuing digital prescriptions, and scheduling appointments. The rest of the time is reported to be spent on anamnesis, physical examination, measurements and tests, and counselling or teaching. Four physicians out of seven recorded as much as possible during visits;

however, a complete recording is impossible due to a lack of time. Five physicians estimate they document around 50-70% of data during the visit. The other two physicians reported only taking notes during the visits and compiling a medical record after the appointment [17].

Research conducted by James Goldenberg investigates the breadth and burden of data collection in clinical practices. He mentions that clinicians are skilled at observation and medical decision-making but struggle to accurately record and measure the activities. The research describes many touchpoints of data collection and operations at different locations in different formats. Goldenberg reports that it is difficult to determine the most meaningful data to collect with all available information. Another problem is that decisions on which metrics should be recorded are often made without consulting the practicing clinicians, resulting in additional frustration and unwillingness to participate in the data collection process. This has also resulted in many financial incentives to solve the data collection problem and has had significant financial repercussions [18].

By now, numerous examples have been given to show the extent of data and data collection problems in healthcare and its various professions. The problem repeats throughout the whole system. Data needs to be collected fast, but simultaneously, it needs to be complete and consistent for future reference. On the other hand, medical professionals should not gather data that will not be needed or used later. Regarding mental health care, a recent study realized that digital platforms could help solve the data collection problem. However, there is a lack of evidence supporting such systems' usability and usefulness. Data collection in mental health care is challenging due to the lack of reliable biomarkers. Therefore, there is also a lack of usable data since a big portion of this data is either on paper or in a free form of text. Due to this, many studies have collected and cleaned digital data for a specific purpose of the same study, proving various benefits of using computerized decision support tools, but not contributing to the lack of usable digital data in mental health care. The research clarifies that more studies are needed to collect more digital data for mental health care [19].

We have established that mental health care has seen some advancements with the help of technology. However, portions of mental health care are still left behind. What happens during the interview sessions between a patient and a psychiatrist either continues to be carried out in the traditional way or has the mental health care provider delve into their devices and not concentrating nor spending enough time with the patient in the room. Most of the innovation in mental health care is done on the patient side – apps that track their lifestyle, help them communicate with psychiatrists, or allow them to self-assess their mental status. Medical professionals have not seen significant advancements in their daily work. A recent study looking into computerized decision aid (CDA) states that despite all the innovative IT solutions, modern tools are rarely used in mental health care [20]. Decision support systems have a central role in health care; however, only a few exist for psychiatric problems. Clinical decision support systems face certain problems, such as the lack of electronic medical records (EMR) due to many technologies not being widely available or used. [21]. The lack of medical records for mental health and the use of handwritten or free text data from psychiatrists have led to studies that usually end up with a dataset within a narrow scope, like schizophrenia, depression, and others. While analysis provides valuable insights to the diagnosis process, they tend to overlook the method of data collection, sometimes even goings far as to clean the data up specifically for their research [22]. Former research comparing the difference in diagnostic outcomes between the traditional pen-and-paper method and a computerized decision support system concluded that there are no major differences between the two, and that traditional decision-making is as effective as computer-supported one. In some demanding cases, clinicians even favoured the pen-and-paper method [23].

To sum up, we can use AI for decision support, but it requires a lot of data to work well, and it takes too much time and effort from professionals to input this data. Data collection is the most important activity in the day-to-day operation of a psychiatrist which is overlooked by many.

1.2.DocuMental

DocuMental is a clinical decision support system (DSS) developed by Documental Ltd in collaboration with various digital health vendors, universities, and healthcare institutions. The system is designed to aid in diagnosing, treating, and managing mental disorders and provides a platform for digitized clinical workflows in mental health. DocuMental is a web-based software tool created for professionals within the mental health service field. These professionals range from physicians, nurses, psychiatrists, social workers, and care coordinators to healthcare managers. The system delivers immediate results, improves user-friendliness, improves clinical workflows, generates documentation, increases time-effectiveness, provides progress overview, increases accuracy, reduces errors, increases confidence, and supports evidence-based decisions [24].

The system has digitized, structured, and automated history and assessment modules for comprehensive personal history, mental status, physical examination, risk assessment, and clinical scales. These modules help clinicians perform comprehensive and standardized assessments for improving disease management and personalized care [24].

The diagnostic module includes structured ICD-10 diagnostic criteria and “tick mark” choice questionnaires for all mental health disorders, supporting adult, adolescent, and child diagnostics. The ICD-10 diagnostic questionnaires are linked to DSS algorithms that provide diagnostic suggestions to increase diagnostic accuracy and allow for diagnostic verification and differentiation [24].

The treatment module includes a complete list of psychotropic medications registered in the European Union for mental health disorders. It operates through DSS algorithms for medication and treatment-plan selection, including doses, regimens, and drug-drug interactions in agreement with widely recognized medical guidelines. The treatment module is linked to the diagnostic module. It tracks medication choices following clinical presentation and diagnosis, which helps clinicians avoid mistreatment and manage care plans in a standardized way [24].

DocuMental is a rare clinical DSS with digitized and structured ICD-10 diagnostic criteria for all mental disorders. It also provides a clinical presentation to treatment registers tracking medication choices and management of care plans in a standardized way.

DocuMental suits a broad range of users involved in mental health services provision. DocuMental also allows patients to contribute to bilateral communication with mental health organizations and professionals [25].

EIT-Health selected DocuMental as one of 21 finalists of Business Plan Aggregator 2016 among digital health innovations and was the only finalist in the mental health sector. It also received an Estonian-Finnish Quality Innovation Award in 2017 [25].

DocuMental is still in an experimental piloting phase and in the process of getting certified. A few psychiatrists use the tool, including the CEO and Founder of the company, Dr. Eduard Maron, who is the leading expert on the system due to his direct involvement in developing the application [26].

This research was done in cooperation with DocuMental.

1.3. Aim and research questions

This Master's Thesis aims to evaluate options for improving the data collection process in psychiatric decision support. The thesis looks to answer the following research questions:

1. How can decision-making algorithms in a decision support system be mined for the theoretical shortest path to diagnoses?
2. What is the actual user journey based on the usage logs?
3. How big is the gap between theoretical minimum and actual usage patterns?
4. What are the main contributors to the inefficiency?

2. Methodology

This section provides detailed information about the study methodology. The author did not do data collection; instead, multiple data sources, such as event logs and data hierarchy and rules, were extracted from the DocuMental application. The extracted data contained multiple files specific to the application, such as the application's event log, the data structures used and concept mappings between various used codes, as the application uses ICD-10 and custom coding of questionnaire answers. The event log used in the research was dated February 21st, 2020, to February 1st, 2021. This study was quantitative, as events were parsed and analysed in sum, and categorical values of these events were analysed in bulk. For privacy protection, the data was extracted in a depersonalized way. The author had no visibility or knowledge of the gender, age, location, or other personal information of the patients.

2.1. Process mining method

The methods used in this research were quantitative, as data was analysed and looked at as a sum, differentiating between categories and various questions, diagnosis, and other data points. Looking at the data from a broader perspective allowed us to have better visibility on actual usage of the checklist in the application and analyse various metrics across different data points. Process mining techniques were used to analyse the usage logs and data structures of DocuMental.

Process mining is a data-driven methodology that aims to discover, monitor, and improve real-world processes by analysing event logs. According to the process mining manifesto, 3 main groups of techniques are used. The first is process discovery, which mainly consists of extracting data and visualising processes models from events. The second is conformance checking which compares the actual model to a predefined one to discover differences and find patterns for improvements. The third group is enchantments, which is about adding additional information to the models for more advanced analytics [27]. However, a recent guide called the Ultimate Process Mining Guide has outlined additional steps within these groups [28]. The current research has determined to use the methodology with additional steps, as defined in the Ultimate Process Mining Guide, for better clarity and division of tasks. The steps are as follows:

1. **Data collection**: At first, it is necessary to gather data about the process of interest by recording events or activities that occur during the process. This data is typically stored in event logs containing information about the events, such as their timestamps, the activities performed, and the resources involved. For this step, the data collection process has already been handled by the DocuMental application and its users. For the study, the data will be extracted from DocuMental database. For privacy protection, the data will be extracted without attributes that could be used to identify a person. Extracted data will be securely transferred to the author.
2. **Pre-processing**: Secondly, it is necessary to clean and pre-process the event logs to remove noise, handle missing data, and transform the data into a suitable format for analysis. For this research, the data extracted in the previous step will be cleaned and processed to a more approachable format, which will be usable in later steps. This step will include data structure visualization that can be used in the next step as well. A program or a script is created for event logs processing. The results, data structures, and visuals created during this step need to be actionable in further research and include some high-level understanding of data structures and containing elements. In addition, the data will be cleaned and filtered using various rules for data exclusion.
3. **Process discovery**: Using process discovery techniques will automatically extract a process model from the event logs. Process discovery algorithms typically generate a graph or diagram showing the flow of activities in the process and any decision points or loops. This step of the research will use data structures and visuals from the previous step to analyse the general usage of the application. In addition, usage patterns and statistical analysis of various metrics will be done.
4. **Conformance checking**: This step involves comparing the process discovery algorithm's process model with the actual event logs to identify any discrepancies or deviations from the expected behaviour. Conformance-checking techniques can help identify bottlenecks, inefficiencies, and other issues in the process. In this step, the research will focus on the actual usage patterns of the psychiatrists. The theoretical shortest path to diagnoses will be compared to the actual paths psychiatrists took. A few use cases will also be described and visualized for the research.

5. **Performance analysis**: Next is to analyse the performance of the process by computing various metrics, such as throughput time, cycle time, and resource utilization. Performance analysis can help identify areas of the process that are performing well and areas that need improvement. During this step of the research, several metrics will be analysed, such as action and question amounts, screen transitions and time statistics for possible improvement areas. For this analysis, a secondary script or program needs to be created to traverse through previously parsed event logs and analyse the metrics from the actual usage data.
6. **Enhancement**: The next step is to use the insights gained from the analysis to improve the process by changing the process design or implementing new tools or technologies. All of the previous steps until this one will be discussed and analysed during the Discussion chapter below to explain possible improvement areas; however, implementing these improvements will not be a part of this research, as the author does not have access to the source code of DocuMental. This and further steps would also increase the magnitude of the current research beyond a Master's thesis limits; thus, these steps should be handled in separate research.
7. **Monitoring**: Continuously monitor the process to ensure it is performing as expected and identify any new issues that may arise. This step will not be used in this research.

Process mining can be applied in various scientific research domains, including healthcare, finance, logistics, and manufacturing. The methodology can be used to study different types of processes, such as clinical pathways, supply chain management, and customer service. The insights gained from process mining can help researchers identify best practices, improve organizational performance, and support evidence-based decision making.

Individual data points will be analysed to structure a pathing of questions answered by the professional, to see which questions have been skipped and if there was any chaotic behaviour – jumping between questions, going back to previous questions on another page etc. For such research, the point of time of the data entry is critical. An analysis of this kind allows the author to compare the minimum theoretical path to a diagnosis to the actual path a psychiatrist took to reach a said diagnosis. Analysing such usage data can give potential insights into bottlenecks and improvement focus points to propose new ways of conducting interviews with patients. The change in the interviewing model allows us to improve the data collection process to reduce the data collection portion of the whole patient-psychiatrist encounter, allowing for a faster diagnosis and leaving more time for the psychiatrist to concentrate on the patient and start the treatment pathway quicker.

The parameters needed for such analysis include:

- The data structures and rules used by the expert system regarding answer options to each question and how they contribute to certain diagnoses.
- And the historical usage data regarding actions taken by the mental health care professional within the application.

Using the previously mentioned data structures, rules, and historical usage data, we can use process mining techniques to analyse the velocity and effectiveness of the current diagnosis checklists.

2.2. Data inclusion-exclusion criteria

The quality of the results obtained from process mining techniques largely depends on the source data quality. In particular, the inclusion and exclusion criteria applied to mental health data can significantly affect the accuracy and reliability of the mining results of the process. Therefore, careful consideration of the criteria of appropriate inclusion-exclusion is essential for achieving an accurate and meaningful understanding of the mining techniques in mental health. The inclusion-exclusion criteria the author has chosen for this research are:

- **Diagnosis checklist** – in medical data, especially in the psychiatric field, where most data is subjective and must be interpreted by the psychiatrist during patient encounters, different questionnaires and diagnoses should not be compared to each other but should be taken by categories or analysed individually. For a better focus in this research, the author has chosen to only analyse encounters with a confirmed anxiety diagnosis further. Anxiety diagnoses were selected to analysis due to the diagnosis checklist being 2nd most used questionnaire, coming right after affective disorders. However, the anxiety data structure was more simplistic and approachable, with also a smaller set of diagnoses being present.

- **Confirmed results** – to clearly understand a psychiatrist's path to reach a specific diagnosis, we can only look at encounters with a confirmed diagnosis.
- **Time aspect** – Only encounters with answers recorded for more than 5 minutes. In addition, as encounters can span over multiple dates, we only look at encounters that have been filled with less than 2 hours. Encounters with more than 2 hours recorded may be valid patient cases; however, long times may reflect another issue with the workflow which the application does not solve.
- **Outliers** – we should also remove outliers by time and actions taken, questions answered, etc. from the data analysed. The outliers are encounters with less than the minimum number of questions answered, or a checklist filled implausibly fast, which is not credible in a real-life scenario.

2.3. Ethical considerations

Data collection processes for patients in psychiatric care pose unique ethical considerations that must be carefully addressed. Firstly, patients' data privacy must be protected throughout the entire data collection process, from the moment the data is collected to its storage, analysis, and sharing. Patients in psychiatric care are vulnerable, and their personal and medical information must be handled with utmost confidentiality to prevent stigmatization, discrimination, and potential harm.

Secondly, the over-collection of data must be avoided, as this could lead to the breach of patients' privacy and loss of trust in the healthcare system. Researchers and clinicians should ensure that they collect only relevant data necessary to improve the patients' care and wellbeing, not intrusive or unnecessary data. Additionally, obtaining informed consent from patients before collecting their data is crucial to ensure transparency, autonomy, and respect for their rights and dignity.

This research does not generate any new clinical knowledge but analyses the current process workflow to optimize data collection. To protect patients' privacy, the data is extracted in an anonymized format from DocuMental database, removing any indicators and connections to actual patients. This data is in a predetermined question-answer format and does not include free-form text, which could be linked to any patient. The data is analysed in the author’s computer, separate from the DocuMental production computers.

There are possibilities for errors in every step of the research. Mistakes could be made by the psychiatrist entering the data into the system. Errors could happen when data is being written to and stored in the database. Additionally, the queries made to the database may include mistakes by the querier or program used to read the data. Errors may have been made when anonymizing the data used in this research or when transferring it. As a single person conducted this research, there may exist errors or logical flaws in the programming codes and scripts written to examine the data structures or during the analysis of data entries. If there are flaws in the processes used during this research, this can affect the outcome of the analysis. However, there are not many decisions that the author has made by themselves. Most of the processing has been automated; thus, a machine calculates the statistics. As a countermeasure, the conclusions made in this research will be used by DocuMental to improve their expert knowledge of result validation. In case third parties wish to carry out similar research, additional methods can be used to validate the correctness of the results.

This research was approved on April 15th, 2021, by the Research Ethics Committee of the National Institute for Health Development. Approval references are research no 2248 and decision no 704.

3. Results

This chapter provides information about the data and processes used in the DocuMental application to reach a diagnosis. This section provides a general overview of the data population, cleaning the data, which criteria were used to exclude outliers, and dives into the discovery process to understand how psychiatrists have been using the application. In addition, time statistics was a focus point to analyse potential improvements to the data collection process during a psychiatric encounter.

3.1. Population and data description

The data provided by DocuMental and used in this research is described in **Table 1** below and includes 11 files:

Description	Format	Size KB	Lines of data	Notes
Mental status questionnaire	JSON	2 579	37 642	Triage questionnaire
Checklist questionnaire bundle	JSON	8 145	115 396	All checklist questions and answers
Diagnosis conditions checklist	JSON	4 487	85 892	
Mental status assessment rule explanation sheet	CSV	450	1222	
Anxiety disorders assessment explanation sheet	CSV	89	416	
Concept mapping DCL-to-DCL (checklist - checklist)	JSON	367	10 090	
Concept Mapping DMS-to-DCL (mental status - checklist)	JSON	1 912	45 863	

Mental status expert rule set	Java Drools Business Rules	763	22 897	2 861 expert rules
Checklists expert rule set	Java Drools Business Rules	2 652	66 021	4 203 expert rules for all checklists
Event log of DocuMental application usage	JSON	1 032 784	75 344 360	58 977 event entries (clicks, answer selections, page switches etc.)

***Table** 1. DocuMental provided data population used in the study.*

The mental status questionnaire represents the DocuMental custom triage questionnaire used to help the psychiatrist reach the correct diagnosis checklist. This JSON file contains the structure of this questionnaire and holds enabling rule data. The mental status questionnaire also holds values for some answer options contributing to specific diagnosis checklists. Once a certain threshold has been achieved, the checklist becomes available for the psychiatrist to continue their work. The JSON file follows the FHIR Questionnaire data structure.

The checklist questionnaire bundle contains the same data as the mental status questionnaire. However, this file holds the data for all other checklists, e.g., anxiety, somatoform, sexual disorders, eating disorders etc.

Mental status and diagnosis conditions checklist file is a JSON formatted file that includes the rule formula for each diagnosis and defines which questions and answer options contribute to that specific diagnosis. The mental status checklist, in some ways, duplicates the data that was also defined in the questionnaire JSON file; however, the specific use cases for this file within the DocuMental application were not disclosed for this research. In addition, this data includes information on which other questions, answer options, and diagnoses are either included or excluded if specific questions or answer options occur. This data is in a custom structure and does not follow any specific FHIR resource.

The mental status and anxiety disorders assessment explanation sheets provide supportive and explanatory documentation for this research. These files break down the question code structure and formulas used to define which card and tab each question belongs to and how each question was connected to a particular diagnosis.

The two concept mappings between DCL-DCL and DMS-DCL files contain checklist items that map different answer options to each other. In other words, a specific answer to one question can already answer a similar question under another question or in another checklist. The DMS file also provides a mapping between the answers of the custom mental health questionnaire and other checklists. This valuable file helps save the psychiatrist time by allowing them to remove questions that have already been answered. This file follows the structure of FHIR ConceptMap resource.

The two Java Drools business rule files contained various rules for the questionnaire and checklists. These rules are used in the application processes to determine which questions must be included or excluded from the available list. In addition, various answer-options-related rules contain information on their impact on different diagnoses. This impact can also have a negating effect, removing unrelated questions and diagnoses from the list of available ones.

The event log of the application usage was a JSON stream formatted log file, where each line was a new event entry in a JSON format representing an action. Each JSON object contained the complete checklist structure with previous answers. This allowed insight into which checklist was being filled, in which order the questions were being answered, and how the psychiatrist moved between checklists. The file was structured in FHIR Questionnaire resource structure, where answer values are given with each new action by the psychiatrist.

3.2. Data collection process

A secondary data collection method was used in this research, as the organization owning and maintaining the DocuMental application has provided the author with data structures and historical event logs generated by the usage of the application. The original records were not accessible to the author directly. The records containing sensitive information were generalized and anonymized for the study.

The data collection process of DocuMental takes place during the encounters between the patient and psychiatrist. The answers are gathered during normal conversational flow, or questions are answered by the patient directly. Traditionally, this process would usually be done using pen and paper; however, as described above, DocuMental has digitized checklists following the ICD-10 psychiatric pathways for diagnosis. DocuMental uses an event-sourced database, storing every event in an HL7-FHIR-compliant resource. This means that with each action, the whole structure of the checklist is saved again as a new version with specified identification to differentiate between patients and encounters. Each version of the structure entry contains more data than the previous version based on the identification. As said, these entries also contain the patient's data and other references, including their name and personal information. As the reference can be pseudonymized, it was possible to decide whether only personal data or the complete reference should be removed during the querying process. For this research, the complete reference and personal data were removed from the queried data, leaving unique identifiers only for the encounters. The database was scanned for suitable resource types for specified dates to export the necessary event logs to identify matching events. Patient references and personal data were stripped from these events, and the remaining events were exported as an event log and forwarded to the author.

As described previously, DocuMental uses the ICD-10 Classification of Mental and Behavioural Disorders handbook to base its data structures and categorization. Recorded diagnostic points can be categorized into segments depending on what is being assessed. For example, if the patient's speech is too loud, aggressive, quiet, or fearful, it may indicate some specific mental illness. However, a problem can arise during the session within the natural conversational flow with the patient, expressed by their close family members, or by examining the patient's behaviour while performing an activity requested by the psychiatrist. Questions and assessments like these can be sorted into various categories:

- Impressions
- Questions
- Examinations
- Observations

DocuMental has created a unique system that corresponds to and is validated against the ICD-10 diagnosis classifications. Each of these categories holds multiple questions, some of which can be automatically enabled or disabled regarding which answer option was chosen for another related question, e.g., impressions regarding speech can already include answers for some questions regarding communication under the observation category or questions regarding reading under the examination category. Most answer options contribute to a specific diagnosis unless any options not generating data are selected. These could be "none," "no answer," "not able to evaluate," etc. At some point, one of the diagnoses will become more likely as more answers refer to it. Thus, a mental illness can be diagnosed by a psychiatrist. **Figure 1** below is an example of the mental health status questionnaire. The figure creates a basic understanding of what the mental status assessment lists can contain and how it could be visualized in a decision tree. The reader must note that natural speech disorders, such as apraxia, cluttering, dysprosody, stuttering etc. exist, and the diagnosis part of the visual has been squashed for easier visualisation of the assessment process. It is also important to note that the figure does not include any examples of internal connections between the answers to various questions.

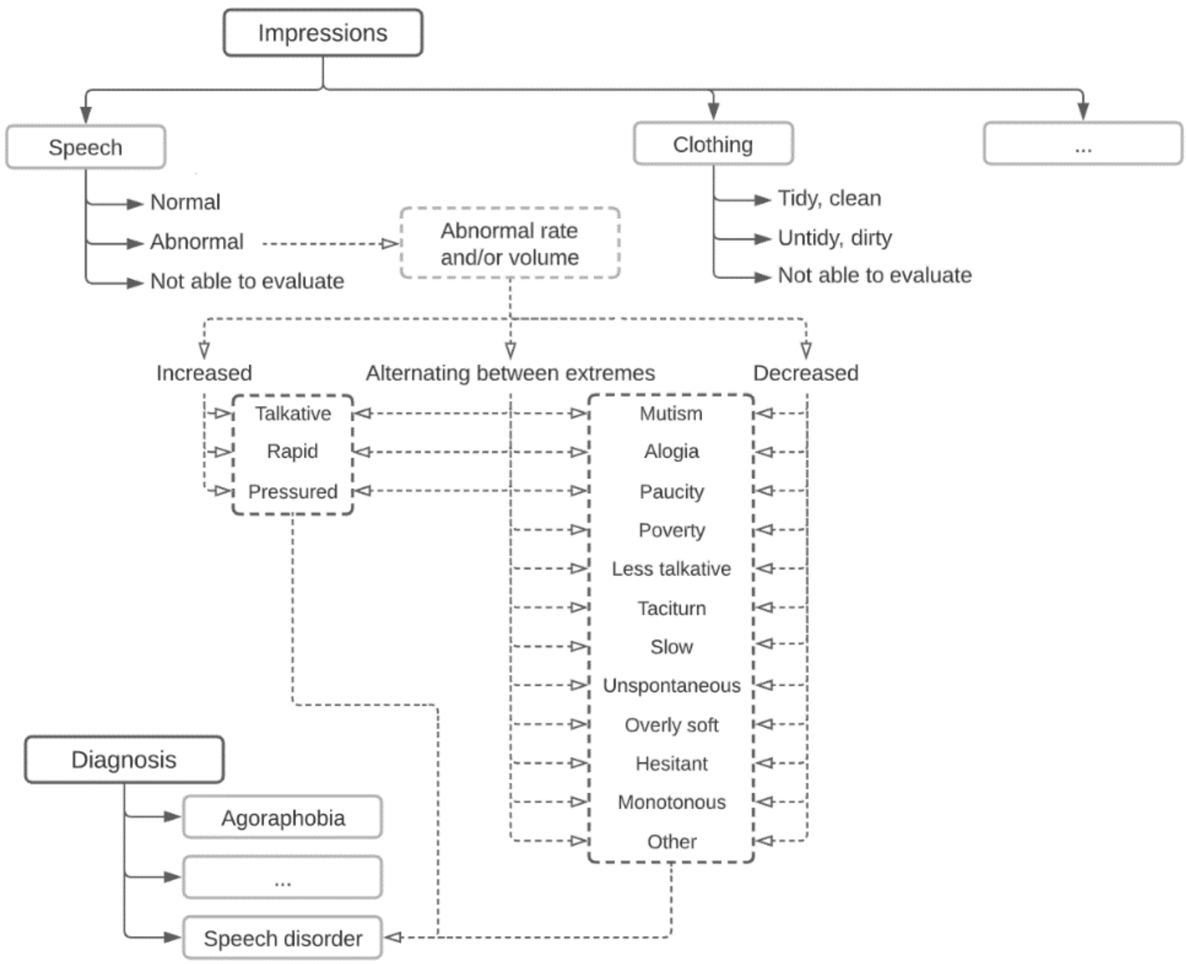

***Figure 1**. Mental status questionnaire example metamodel*

To better understand this figure, an example can be made. During an encounter with a psychiatrist, the patient has abnormal speech. This is being noticed by the psychiatrist and is considered to fall under the impressions category. The patient talks little; their speech is monotonous, without a rhythm, and with an uncommon intonation. The speech has an atypical start and stop pattern. All these impressions contribute to diagnosing someone with a speech disorder like dysprosody. Dysprosody is a common characteristic in some types of aphasia, which is an acquired language disorder usually caused by a stroke or a brain injury [29].

3.3. Pre-processing data

The pre-processing step of process mining was initiated after event logs and data structures were forwarded to the author. This step aimed to clean and pre-process the event logs to remove noise, handle missing data and transform the logs into a suitable format for further analysis.

The data structures of checklist questionnaires were parsed to a more convenient format for analysis using C# programming language as it allows for strict data types. It is an object-oriented programming language that makes it more suitable for handling FHIR-specific data structures. The parsing aimed to create a graph of the resources and to visualize the interconnections between those resources. The created program traversed through the checklist questionnaire resources and created a list of nodes, which included three different types of nodes. The nodes were divided into *Diagnosis*, *Question,* and *Answer* nodes. The checklist questionnaire resource provided a list of questions and answer options for each question. Additionally, the diagnostic algorithms were traversed to create a list of *Diagnosis* nodes, as this specific data was not present in the checklist questionnaire dataset. The program differentiated between 364 diagnosis nodes, 889 question nodes, and 3414 different answers.

As the checklist questionnaire data structures contain references of which answers belong together with which questions, and the diagnosis conditions checklist contains information on which questions contribute to specific diagnoses, the two structures were additionally scanned in parallel for a list of links between said nodes. This allowed the research to preserve the connection of mapped nodes, and 21 749 links across all nodes were identified.

The list of nodes and links between nodes were used to visualize a simple graph to understand better how interconnected these resources are. **Figure 2** below shows a graph for anxiety-related diagnoses, questions, and answers. The large green circles represent a diagnosis. The blue circles represent questions. Orange circles represent answer options for each question. The links between these nodes show how they contribute to specific diagnoses in general. In this graph, we see that specific questions and answers contribute to more diagnoses than others, and some did not contribute to any diagnoses at all. These answer options could also be labelled "not able to evaluate," etc., so these can be legitimate answer options in the graph that do not contribute to any diagnosis.

To further explain this figure, some unused answer options that do not contribute to any diagnoses are indicated by point **A**. Other answer options related to the same question (connected to the same blue circle as the unused questions in point **A**) contribute to a specific diagnosis, as point **B** shows. As indicated by point **C**, some answer options also contribute to numerous diagnoses simultaneously.

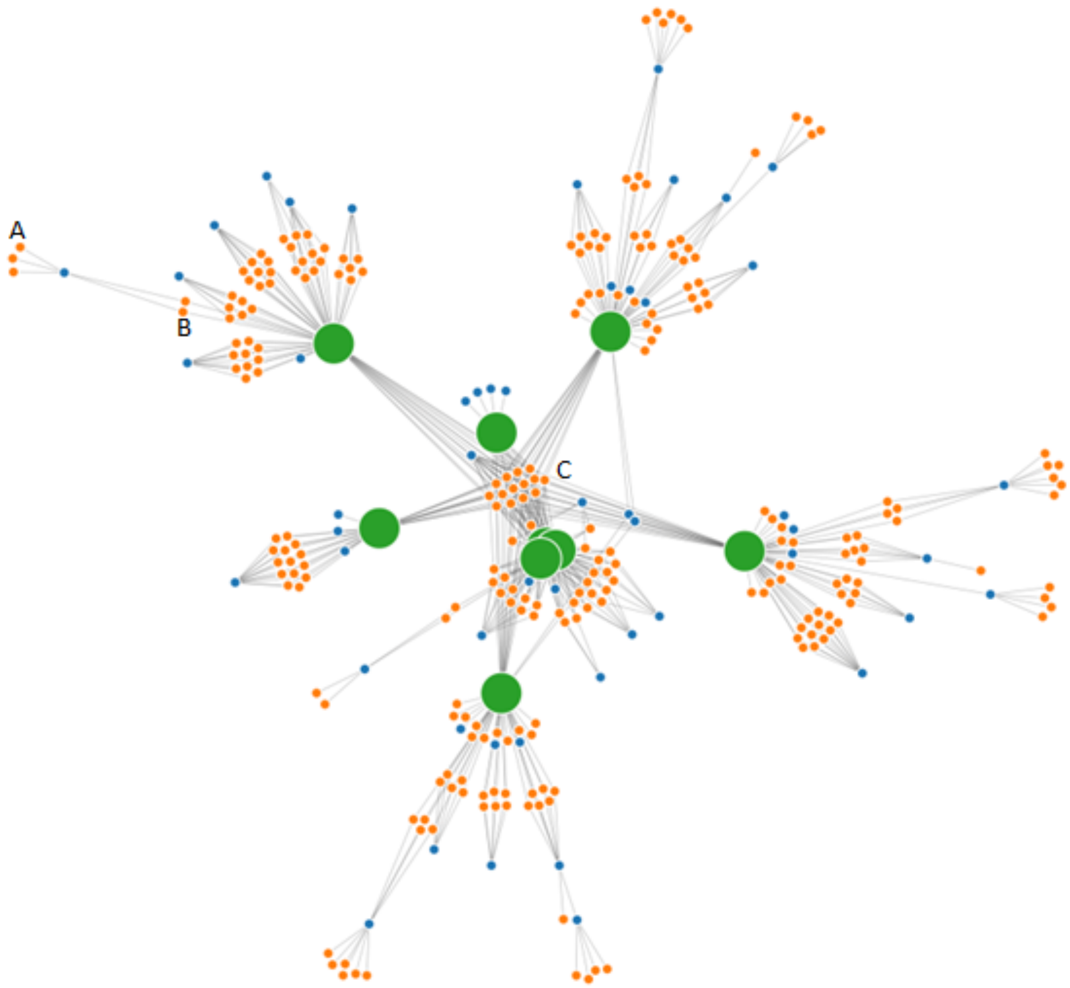

***Figure 2**. Anxiety diagnosis graph.*

The research discovered several questions, and their connected answer options did not contribute to any diagnosis. The existence of these unused questions and answers remains uncertain; however, the DocuMental team has explained that these questions may be in use in the web application itself, separate from the backend logic of the application [26]. These are not present in the following figure.

Additionally, the two concept map files were parsed and traversed to map links between checklists, and additional links were created between nodes.

Furthermore, the expert rule set files were scanned using regular expressions to determine how various questions, answers, and diagnoses were affecting each other, as some answer options could disable other questions or refute specific diagnoses. This information created additional links between nodes, then marked with separate colours to visualize the inclusion and exclusion connections between specific resources. The graph above was a visualization of the process mining results for data structures, which does not include the interconnections between different questions or answers based on rules and concept mappings. This is due to the graph becoming unreadable; thus, a more simplistic view has been made for the reader. This graph is not required for following statistical analysis in this research.

The research continued to investigate the event log files. This file included the complete structure of the handled checklist in each row, whereas the row was a complete JSON object. The rows can be grouped per encounter identifier. On the other hand, each row was a continuation of the encounter with an additional answer option added to the structure. This structure could also hold the information of a newly assigned diagnosis and whether it was confirmed, unconfirmed, or in a single case – refuted. The description of this data structure can be found in **Appendix 1** - JSON stream data structure.

Traversing through the event log and comparing each row against the previous row with a matching encounter ID returned the difference between the two objects. Based on the comparison, a new action done by the psychiatrist was detected. These actions were then separated into a structure where a new file was created for each encounter identifier. This new file encompassed a JSON object holding a list of recorded actions done by the psychiatrist. The list of actions gives an overview of the psychiatrist’s navigation. The action entry included the checklist used as well. Additionally, as the event log rows also included data on assigned diagnoses, an additional list was created to hold all diagnoses and their status for each encounter along with necessary metadata, such as the timestamp of the diagnosis.

3.4. Excluded data

During the research, anomalous data was encountered multiple times, which was why multiple exclusion criteria were used. The exclusion criteria were decided upon during the research. However, looking at this data to validate the excluded anomalies is still imperative.

A single case was still marked as "in progress" in the data. Additionally, a single case of a diagnosis being refuted, either from a confirmed or unconfirmed diagnosis, existed. Still, it cannot be confirmed if this was an intended result due to missing data. There were also two encounters with a diagnosis entry, but no answers were in the dataset. In this case, there was either a possibility that the answers were not extracted from the source system, only the diagnosis was recorded; or simply some test data was recorded.

During the general analysis of the data present, it was found that 39 encounters occurred in more than a single day. The maximum number of days an encounter took place was 39. These specific encounters were marked as outliers and discarded from the rest of the analysis according to the exclusion rules. In addition to that, the DocuMental team notified the author that the application is still in the prototype phase. Due to that, it is being used during live demos and ad-hoc testing during development [26]. Because of these reasons, it was impossible to distinguish between test data and actual data filled by a psychiatrist, at least not on the data provided for this study. This was a significant factor in the data quality and was another main factor in excluding some data from the study. Outliers were detected based on various metrics analysed and described in the next chapter.

Out of 814 encounters, it was discovered that 262 encounters had a diagnosis assigned at some point, whether confirmed or unconfirmed. The number of confirmed diagnoses was 201 across all encounters. Considering the previously set exclusion criteria, **Table 2** below describes the statistics of excluded encounters for anxiety-related data, which directed the research to analyse 39 encounters with a confirmed anxiety diagnosis. These encounters included 7130 actions that were further explored in the next chapter. These criteria were applied in parallel, meaning that, for example, the same encounter might have been an outlier and excluded due to either being out of time bounds or not be related to the anxiety checklist. It is essential to note in the table below that the detected outliers (marked by *) were not associated with any encounters with an anxiety diagnosis. However, these outliers were still excluded from the general population of data. In this case, the outliers would also be a part of the set unrelated to any anxiety diagnosis.

No	Criteria	Number of encounters
1	**Not related to anxiety diagnosis**	771
2	**Filled for less than 5 minutes**	3
3	**Filled for more than 2 hours**	1
4	**Outlier**	13*
	Total excluded	775
	Remaining	39

***Table 2**. Excluded number of encounters.*

Across the whole data population, 151 encounters were analysed further in the general analysis as they were not excluded by the previous criteria, excluding the first criteria.

3.5. Process discovery

Process discovery is a significant component in the process mining technique that involves identifying and analysing business processes from event logs. Process discovery aims to gain insights into how processes occur rather than how they are assumed to occur. By analysing event logs, process mining algorithms can identify patterns of underlying processes, identifying bottlenecks, inefficiencies, and other areas for improvement. Process discovery includes several steps: data extraction, cleaning, process modelling, and analysis. The research will now analyse the remaining data as the data has been pre-processed and modelled, and some data has been excluded.

The event log was parsed programmatically into smaller subsections per encounter, which allowed the research to analyse answering patterns and detect days and times between answers. Statistical analysis was done across this various data in an aggregated manner. During the research, the author created a Python script to analyse the data entries created in the pre-processing step.

Before delving into the parsed data, it is also imperative to look at the full population and general overview of the data. The event log data included 814 patient encounters over the course of 21.02.2020 to 01.02.2021 date range. Within these encounters, 58 978 actions of the psychiatrist were recorded within the application. The research focused only on 39 encounters with an anxiety-related diagnosis, as described in the previous chapter.

Across the data population with confirmed diagnoses, the minimum number of actions that led to a confirmed diagnosis by the psychiatrist in a single encounter was 23, and the maximum was 382, as visualized in **Figure 3** below. The 75th percentile on this data was 187, meaning that 75% of encounters have less than 187 answers given to reach a confirmed diagnosis.

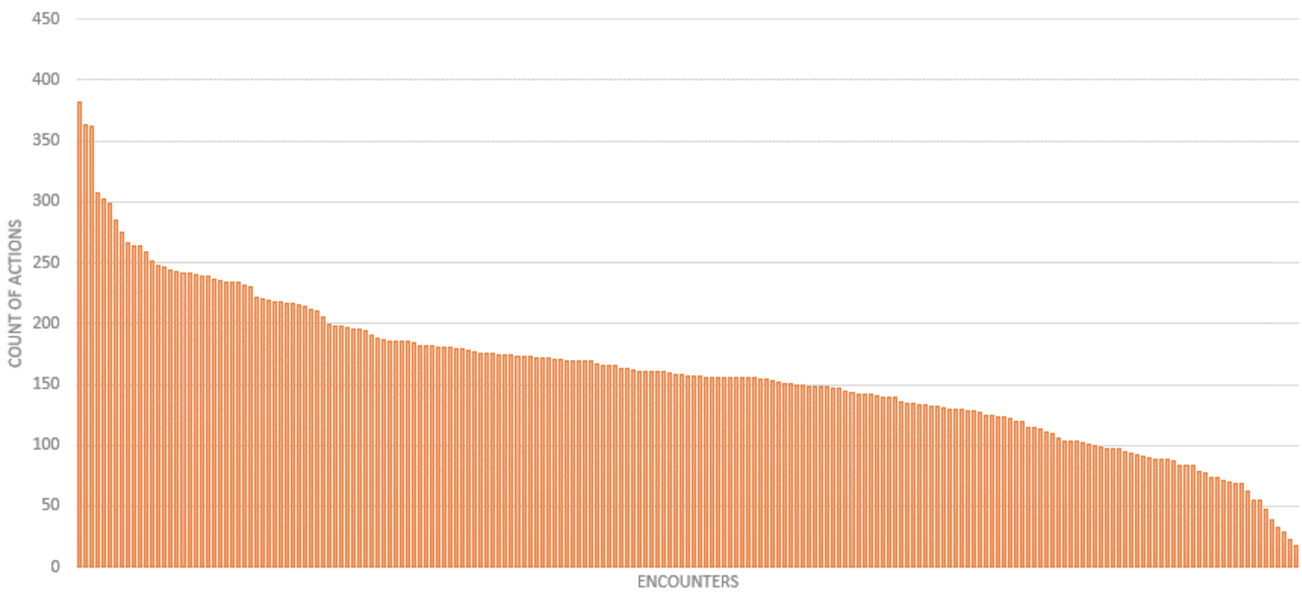

Figure 3*. Count of actions until a confirmed diagnosis per encounter.*

The most extended number of days for a confirmed diagnosis encounter to be in progress was 137, which seemed unlikely considering it was a clear outlier amongst all the other encounters. This encounter was filled for 137 days and had many entries where only a few questions were answered within a single day. This specific entry seemed more like a test instance than an actual patient; thus, the encounter ID was included in the list of outliers for further exclusion. This sort of data point reinforced the usage of exclusion criteria. Of all remaining 200 encounters with a confirmed diagnosis 189 occurred on a single day, and 11 occurred over more than one day. The most extended number of days to fill out a single encounter with a confirmed diagnosis was 20. The average number of days across all encounters was 1,4.

As the questions were answered across multiple checklists, we should also consider the psychiatrist switching between these checklists and between the pages of these checklists. The minimum number of questions answered before switching to another checklist was just one, and the maximum number answered before switching to another checklist was 287. However, the encounter with the most answers in a checklist was again the same, already marked as an outlier. Overall, there were 19 different checklists for various mental assessments. The average number of questions answered before switching to another checklist was 33, and the average number of checklists used across all encounters was 2,56.

With the previously set exclusion criteria, 151 encounters remained. **Table 3** below shows the amount each checklist used across all confirmed encounters. It is important to note that many encounters used multiple checklists; thus, the sum of encounters across the checklists in the table below is higher than the actual number of encounters. 151 encounters with a confirmed diagnosis were still analysed. Further breakdown of these encounters can be found in **Appendix 2** - Checklist usage across encounters.

No	Checklist name	Encounters used in
1	**Mental Status**	150
2	**Affective Disorders**	100
3	**Anxiety Disorders**	70
4	**Personality Disorders**	9
5	**Psychoactive Substance Use Disorders**	8
6	**Somatoform**	8
7	**OCD**	7
8	**Eating Disorders**	6

9	**Psychotic Disorders**	6
10	**Self-Report**	6
11	**Sleep Disorders**	5
12	**ADHD**	3
13	**Tics & Other**	2
14	**Global & Specific**	2
15	**Cognitive Disorders**	2
16	**Dissociative**	1
17	**Stress Disorders**	1
18	**Autism**	1
19	**Social Functioning**	0

Table 3*. Checklist usage across encounters with a confirmed diagnosis.*

For a better focus in this research, only encounters with a confirmed anxiety diagnosis were analysed further. This data included 43 encounters with a confirmed diagnosis. However, based on exclusion criteria, 39 of these encounters were analysed. Various metrics analysed within the encounters with an anxiety-related diagnosis are visualized in **Table 4** below.

Metric	Min	Max	Mean
Duration (seconds)	448,56	2277,75	1033,86
Actions to diagnosis	71	298	167,64
Questions to diagnosis from start of encounter	20	146	94
Questions to diagnosis within anxiety checklist	10	49	21,44
Time to reach diagnosis (seconds)	448,56	1516,76	951,32
Average time between actions (seconds)	4,57	17,79	9,36
Average time between questions	5,1	20,87	11,08

Time to reach anxiety checklist (seconds)	0	1404,07	775,34
Screen transitions	5	20	11,28
Actions after diagnosis confirmation	0	98	13,92
Questions after diagnosis confirmation	0	25	5,23
Time spent after diagnosis confirmation (seconds)	0	806,78	87,04

***Table 4**. Metrics analysed on encounters with an anxiety-related diagnosis.*

The source data set also included two concept mapping files between the mental status checklist and other checklists and different checklist answer mapping between each other. This data was used to check each new action during an encounter to see if the question has any mappings between other checklists and then checked if any of the previous answers were present in the list of equivalent answer options. Simply put, the research looked at data duplication, as different answer options for questions were equivalent and recorded the same data. The research found that this happened on 9232 occasions across all encounters and on 1284 occasions on anxiety diagnosis-related encounters. The arithmetic mean of answer duplication across all encounters was 57,89 actions and 32,92 on anxiety-related encounters. As the mean time between actions was 9,36 seconds, these duplications could be translated to 541,85 seconds lost in general across all encounters or 308,13 seconds within anxiety-related encounters.

3.6. Conformance checking

As the research already pre-processed data to a better format to analyse the actions, the data suffices to create an actual diagnosis pathway the psychiatrist has taken. Additionally, the previously parsed checklists of nodes and links allowed us to analyse the newly created structures to see which questions and answers are necessary to reach a specific diagnosis. The research looked at how experts navigated through the checklists and compared it to the theoretical shortest path to a diagnosis. **Figure 4** and **Figure 5** visualize two cases of the anxiety checklist answer sequence and how the psychiatrist has entered answers to questions. These figures do not show the pathway before the anxiety checklist or other checklists answered after. In these figures, each column represents a single question, and a new row was created for each question that was answered by the professional. The black lines represent the given answer by the psychiatrist, whereas the red line shows the shortest path to confirm the same diagnosis the psychiatrist confirmed.

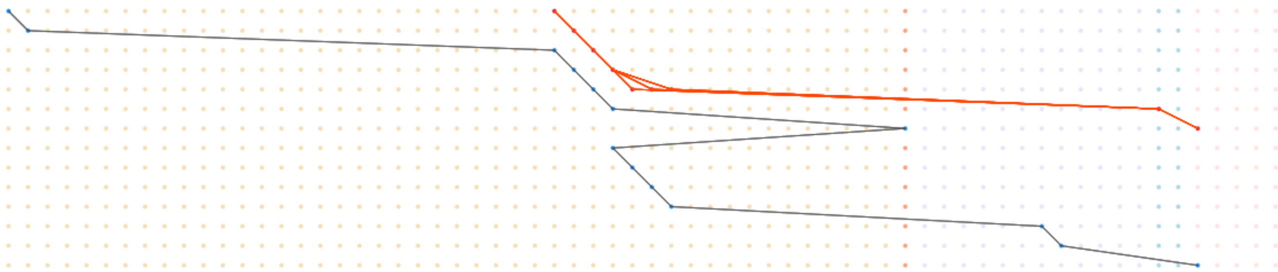

Figure 4. *Visualization of anxiety checklist answer sequence - case 1.*

In the previous figure, we can see that the psychiatrist started filling out the checklist at the very beginning of the questionnaire but switched to a specific question that matches the shortest path of the ending diagnosis. However, during these entries, the specialist jumps to other questions unrelated to the final diagnosis. This specific case was diagnosed with a moderate panic disorder (ICD-10 code F41.00) and had 116 questions answered with 231 actions to reach a diagnosis. The encounter data was recorded for 1225 seconds. The anxiety checklist was reached after 183 actions, 48 actions were done on 14 questions within the anxiety checklist, and 58 actions on 21 questions were done after the confirmed anxiety diagnosis. In psychiatry, as many answers must be interpreted during a conversation, the professional can naturally doubt the diagnosis and approach the patient's worries from different angles. The specialist seems to have started to jump between different questions until he returns twice to the last remaining question for the diagnosis. Even with a confirmed diagnosis, 21 extra questions are answered.

Figure 5. *Visualization of anxiety checklist answer sequence - case 2.*

In the second case, the psychiatrist started filling out the anxiety checklist at a specific spot, corresponding to the start of the shortest path for a particular diagnosis. In this case, the confirmed diagnosis was generalized anxiety disorder (ICD-10 code F41.1). The psychiatrist answered 100 questions over 149 actions. The encounter lasted 1514,86 seconds. The anxiety checklist was reached after 102 actions, 47 actions were done within the anxiety checklist, and the encounter ended with a confirmed diagnosis. No actions were recorded after the diagnosis confirmation.

Comparing the theoretical shortest path to a diagnosis against the actual path taken by a psychiatrist allows us to analyse the usage patterns of psychiatrists to find potential improvement points in the current process. The key aspect of such comparison was to break down and consolidate both data structures for an ideal comparison.

3.7. Performance analysis

The research focused on anxiety checklist and diagnoses connected to this area. Thirty-nine encounters with a confirmed anxiety diagnosis were analysed, excluding data that did not match the selection criteria. The arithmetic mean of questions answered by a psychiatrist for a confirmed anxiety-related diagnosis using the anxiety checklist was 21,44. Anxiety-related diagnosis rules were analysed to create a mapping for the theoretical shortest necessary path to confirm a diagnosis. All diagnoses and their specific question numbers are described in **Table** 5 below. There are cases where additional questions have been answered after the minimal shortest path for a confirmed diagnosis has passed. The arithmetic mean of answers given after a confirmed diagnosis was 13,92 for anxiety diagnosis-related encounters.

Diagnosis	ICD-10 code	Theoretical shortest path questions	Number of encounters with confirmed diagnosis	Number of questions (all checklists / anxiety checklist)		
				Min	Max	Arithmetic mean
Agoraphobia	F40.0	9	4	90 / 13	167 / 48	126,75 / 33,5
Social phobia	F40.1	10	6	53 / 14	132 / 21	94 / 18,5
Specific (isolated) phobias	F40.2	9	1	97 / 18	97 / 18	97 / 18

Panic disorder - episodic paroxysmal anxiety	F41.0	6	14	25 / 10	140 / 49	87,43 / 17,71
Panic disorder - moderate	F41.00	7	4	103 / 14	129 / 40	117,25 / 24,5
Panic disorder - severe	F41.01	7	1	126 / 126	126 / 126	126 / 126
Generalized anxiety disorder	F41.1	5	9	47 / 11	134 / 40	98,11 / 20,89
Depersonalization-derealization syndrome	F48.1	4	0	-	-	-
Anxiety disorder due to known physiological condition	F06.4	4	0	-	-	-

***Table** 5. Anxiety-related diagnosis theoretical paths compared to actual number of questions answered.*

When comparing the actual average arithmetic mean of questions answered by a psychiatrist to reach a diagnosis within the anxiety checklist, and the theoretical shortest paths, there exists an average difference of 14,66 questions, which could be used to reduce the number of actions and time taken overall. Considering the average time between questions was calculated 11,08 seconds, the theoretical and actual question amount difference is roughly 162,43 seconds. However, due to the nature of the work of psychiatrists, it is not possible to say for sure what the improvements could be, as a significant factor in the process is what happens during the encounters with patients and what is being discussed with them, as there may be other contributing factors that take additional time to answer specific questions.

Additional findings were found during the process discovery step as well. The research discovered that there existed 188 question nodes that were not connected to any diagnoses. The unused questions were cross-checked and confirmed with the rule files and Excel worksheets. Fifteen unused questions existed in the anxiety diagnosis structure. These questions were used on 1286 occasions. In addition, these questions were used an average of 5,67 times per encounter.

3.8. Time aspect

Questionnaires are a common method for gathering information about mental health symptoms and patient outcomes in psychiatric practice. Sometimes, the same questionnaires are still being filled using traditional pen-and-paper methods. By analysing the time-related patterns of questionnaire answers, psychiatrists can gain valuable insights into patient symptoms and treatment outcomes, leading to improved clinical decision-making and workflow optimization. Time-related data log analysis techniques offer insights into data collection methods, bottlenecks, and shortfalls. The main contributors timewise during an encounter are real-life situations and conversation, application load times, answer option selection by the psychiatrist, and navigation within the application - screen transitions between different checklists, questions, and answers. This research could not evaluate the time various activities take during an encounter and the application load times. Hence the research will focus on the actions done within the application by the psychiatrist and the screen transitions while filling out questionnaires.

As patients can have multiple sessions with the same encounter identificator on different days with the psychiatrist, tracking the time and sessions it took before a diagnosis was reached was also possible. As mentioned in the Process discovery chapter, the minimum number of days/sessions to reach a diagnosis was 1, and the most prolonged period of days/sessions a psychiatrist has spent before diagnosis was 20. The average number of days to reach a diagnosis was 1,4 across all encounters. The research excluded all encounters which took less than 5 minutes or more than 2 hours to fill. Within those timeframes, the least time an encounter took was 304,07 seconds, the most was 7127,95 seconds, and the arithmetic mean across 151 encounters was 1059,43. The anxiety-related time statistics can be found in **Table 4** above. The arithmetic mean across all anxiety-related encounters was 1033,86 seconds, and the arithmetic mean time between different questions was 11,08 seconds.

The question-answering sequence was analysed to understand screen transition amounts the psychiatrist has made while using the application. As questions within checklists are divided between different tabs, the research counted screen transitions between tabs marked within question codes. The most screen transitions a psychiatrist had made during an encounter was 22, and the least was 4. The minimum number of screen transitions for anxiety-related diagnoses was 5, and the maximum was 20. The arithmetic mean for screen transition across all encounters was 11,07 and 11,28 for anxiety-related diagnoses. As DocuMental is a web-based application, we should consider website load times. The average time to load a webpage is 10.3 seconds [30], meaning a screen transition with the psychiatrist orienting to the new page or questions displayed could be around 11 seconds per transition [31]. The arithmetic mean for potential screen transition time-saving was 124,08 seconds, which could be an 11,72% improvement to the 1059,43-second encounter duration. Additionally, the research compared the average time spent between all questions to the average time spent on two questions on different tabs. The time difference between these two metrics was 18,06 seconds, which was more than the suggested time by Brian Dean. The research proceeded with the time indicated by Brian Dean with a small addition of orientation time, amounting to 11 seconds.

Within the analysed 39 anxiety-related encounters, the arithmetic mean time for a psychiatrist to reach the anxiety checklist was 775,34 seconds. The mean time it took a psychiatrist to reach a diagnosis within the anxiety checklist was 175,98 seconds, averaging 71,41 actions. The overall arithmetic mean time and actions for a psychiatrist to reach a diagnosis from the start of the encounter was 1033,86 seconds or 167,64 actions across all checklists. In addition to screen transitions and navigating through checklists, additional time is spent on questions and actions after the diagnosis confirmation. The arithmetic mean time spent after confirmed diagnosis was 87,04 seconds.

The encounter with the maximum number of 22 screen transitions contained 29 actions after the confirmed diagnosis and 414,13 seconds spent on those actions. The assumed screen transition time for this encounter could be 242 seconds (226,6 load time + 15 seconds for perception), the sum of these times could amount to 656,13 seconds. As the diagnosis in said encounter was reached in 860,61 seconds and the whole encounter took place in 1274,74 seconds, the potential time-saving in this specific case could have been 51,47%. **Table** 6 below gives an overview of the time-related metrics for anxiety-related encounters. The reader should note that as shown in **Table** 5, some of these diagnoses only have a single case with a confirmed diagnosis, hence why the minimum, maximum and mean are identical.

Diagnosis	ICD-10 code	Time to reach diagnosis in seconds		
		Min	Max	Arithmetic mean
Agoraphobia	F40.0	978,14	1101,11	1040,64
Social phobia	F40.1	773,17	1465,42	1059,7
Specific (isolated) phobias	F40.2	1505,9	1505,9	1505,9
Panic disorder (episodic paroxysmal anxiety)	F41.0	448,56	1510,07	869,06
Panic disorder - moderate	F41.00	576,24	1490,01	980,04
Panic disorder - severe	F41.01	855,82	855,82	855,82
Generalized anxiety disorder	F41.1	481,22	1516,76	903,57
Depersonalization-derealization syndrome	F48.1	-	-	-
Anxiety disorder due to known physiological condition	F06.4	-	-	-

***Table** 6. Diagnosis related time statistics.*

In addition, **Table 7** below describes various metrics on lost time.

Diagnosis	ICD-10 code	Mean screen transitions	Screen transition potential lost time (seconds)	Mean time spent after diagnosis confirmation (seconds)
Agoraphobia	F40.0	12,25	134,75	221,5
Social phobia	F40.1	12,83	141,17	84,06
Specific (isolated) phobias	F40.2	13	143	174,19
Panic disorder (episodic paroxysmal anxiety)	F41.0	9,93	109,21	50,11
Panic disorder - moderate	F41.00	13,75	151,25	245,64
Panic disorder - severe	F41.01	13	143	0
Generalized anxiety disorder	F41.1	10,44	114,89	16,22
Depersonalization-derealization syndrome	F48.1	-	-	-
Anxiety disorder due to known physiological condition	F06.4	-	-	-

***Table** 7. Diagnosis metrics for potential time-saving.*

The arithmetic mean of time spent after a confirmed diagnosis was 87,04 seconds within the anxiety checklist, which could potentially be entirely removed. As there also existed unused questions, the research should also account for those. The mean time between answering questions was 11,08 seconds; thus, the time spent on unused questions per encounter was 62,82 seconds. The discovered potential improvements of reducing question amounts, removing data duplication and unused questions, reducing screen transitions, and eliminating time spent after a confirmed diagnosis sums up to 744,5 seconds. As the arithmetic mean duration for 151 encounters was 1059,43 seconds, the percentage of potential time-saving is 70,27% per encounter.

4. Discussion

This section discusses previous research results and tries to answer study questions. Moreover, the main contributors of this study are upbrought, and the limitations of this study are examined. Additionally, perspectives for future research are given.

4.1. Input data and collection

Data capturing is essential to any medical diagnosis, serving as the foundation for deriving insights and making informed decisions. The process of capturing data in medical fields, however, can be a complex and time-consuming task, requiring the careful judgement of various external factors, technical limitations, and considerations about which type of data needs to be collected, the methods used to collect it, and the tools and techniques operated to analyse it. As such, optimizing the data-capturing process is crucial to ensure that medical care is carried out efficiently and accurately and that real physical and mental health issues are addressed as soon as possible. This section discusses the steps taken to optimize data capture in a mental health care application using process mining. This methodology leverages data from existing systems to analyse and improve processes. By analysing data from the research phase, we can identify bottlenecks and inefficiencies in the data-capturing process and implement measures to streamline the process and enhance data quality. This topic was researched as mental health issues are becoming more open in today's society, as well as the growing popularity of mental health applications. From the defined steps of process mining in chapter 2.1, this research did not improve any processes nor monitor the processes of DocuMental for these improvements. Thus points 5 and 6 of previously defined process mining steps were omitted.

The research was conducted by analysing data from a mental health digital decision support tool, DocuMental, developed and used by a few psychiatrists in Estonia. It is also important to note that the application is still in an experimental piloting phase and in the process of getting certified [26]. The data spans 346 days, having data for 814 patient encounters with psychiatrists. The data analysed includes 58 978 recorded actions within the application.

4.2. Pre-processing data

The data provided by the DocuMental team was of good quality and in a structured format, allowing for an easy approach to analysis. Most of the data provided was FHIR compliant, allowing additional documentation to be used for data structure understanding and parsing. Extra Microsoft Excel and comma-separated value files were provided with an explanation for understanding. These provided a good overview of interpreting the DocuMental custom formulas for diagnosis confirmation. As most of the files were in a suitable format and mainly described diagnoses, questions, answer options, their connections to each other, and how certain answer options contribute to specific checklists and diagnosis confirmations, these files were mainly used to understand the structure of the processes and did not need any additional cleaning. These structures were parsed into a simpler graph structure of nodes and links, which were then used to visualize the checklist structures containing questions, answers, and diagnoses. An example of an anxiety checklist structure visual was given previously in **Figure 2**.

A substantial part of the data cleaning process went to processing the given event log to a better readable format where each action made by the professional can be analysed separately. An algorithm was created using the C# programming language, as it allows for strict data structures, which simplified the processing of event log file into smaller portions. The data was divided into separate files per encounter and actions done within an encounter. These actions mainly contained chosen answer options but included the name of the checklist, question category (i.e., tab), and a timestamp, which could later be used in the process discovery portion of the research.

4.3. Data exclusion

The research focused only on encounters with a confirmed diagnosis result that the psychiatrist assigned. Furthermore, only anxiety diagnoses were analysed in detail. Data entries were excluded by the various criteria noted in chapter 2.2. Among the previously mentioned 814 encounters, only 262 had a diagnosis assigned at some point. The rest 552 encounters had no diagnosis or outcome within the data timeframe. In addition, not all of these diagnoses assigned to encounters were confirmed. The number of confirmed diagnoses analysed was 151 across all checklists and encounters; however, one of these encounters was an outlier and excluded from further analysis. 150 encounters with a confirmed diagnosis were analysed for general results regarding actual paths to diagnosis by psychiatrists and time analysis. When considering all previously stated exclusion criteria, 39 encounters had a confirmed anxiety-related diagnosis and were used as the main subject of this research. These encounters contained 7130 actions which were analysed in-depth.

The research discovered outliers, where certain encounters did not include any data and encounters with diagnoses assigned to them. However, data was absent for these encounters. Additionally, numerous encounters took place over more than a single day, which were excluded. Furthermore, the research excluded encounters where a diagnosis was assigned within 5 minutes from the start of the encounter due to DocuMental being used in live demos [26]. Outliers in the source data were the sole purpose of creating exclusion criteria to improve the data quality.

4.4. Process discovery

The research continued in the process discovery phase, which involved analysing the event logs of the DocuMental application and its usage by psychiatrists. The question structures, their division into different tabs or pages, and the pre-processed event log answer sequences were combined to visualise the path of actions taken and questions answered by a psychiatrist to reach various diagnoses in the anxiety checklist. From there, the research saw initial inefficiencies and potential for further analysis.

The pre-processed log analysis continued via a Python script that analysed all encounters and actions. The data was split into multiple portions - diagnosis-specific data was consolidated in a data structure that consisted of summed-up statistics for each diagnosis. This data for each diagnosis included various metrics, such as encounter counts, encounter durations, actions made overall, actions made within the anxiety checklist, questions answered, questions answered within the anxiety checklist, time to reach the anxiety checklist, time spent in anxiety checklist, time to reach a diagnosis, time spent after diagnosis, actions made after diagnosis, questions answered after diagnosis, screen transition counts and transition times. These metrics have been summed up to gain insights into anxiety-specific diagnoses and their statistics. The processing results are available in **Table 3** and **Table 4** above. When comparing the average arithmetic mean of questions answered by a psychiatrist within the anxiety checklist and the theoretical shortest paths to reach an anxiety-related diagnosis, there exists an average difference of 14,66 questions, which could be used to reduce the number of actions and time taken overall.

4.5. Conformance checking

Currently, DocuMental application works as a list of questions with simple suggestions for checklists to use during the session or when a diagnosis has enough data to confirm. The questions within a checklist can exist on different pages depending on the groupings defined for the checklist. The data structures were parsed to a more actionable format which contained encounters and the psychiatrist-made action sequence within the encounter. The research visualized two cases where a comparison between the actual question sequence of the psychiatrist and the theoretical shortest path to the same diagnosis could be seen in **Figure 4** and **Figure 5**. This point of process mining noticed excess questions being answered in the anxiety checklist.

4.6. Performance analysis

The general analysis of the application usage data determined that the overall arithmetic mean of actions taken by a psychiatrist for a confirmed anxiety diagnosis was 167,64. The mean number of actions a psychiatrist took for a confirmed anxiety-related diagnosis using the anxiety checklist was 71,41. When comparing the actual number of actions taken by a psychiatrist and the theoretical shortest paths to reach a diagnosis within the anxiety checklist, there existed an arithmetic mean difference of 14,66 questions, which translates to roughly 162,43-second time-saving potential. This shows the possibility for improvement by reducing the number of actions and time taken overall. As said before, the work of a psychiatrist largely depends on the patient and conversation flow. It should be analysed further by researching and testing out various methods on how the application works and how it could be used during encounters in an efficient manner.

4.7. Time aspect

The research proceeded by analysing time-related statistics of the data set. Within the 39 encounters with a confirmed anxiety diagnosis, it was determined that the mean time for the psychiatrist to reach the anxiety checklist was 775,34 seconds and that it took 951,32 seconds for a psychiatrist to reach a diagnosis with a mean of 167,64 actions.

Based on the statistical analysis of the final filtered data, the research compared the general population statistics to the statistics of the anxiety subset. It confirmed that the amount of data analysed represented the general population data well. It was sufficient to provide a broad understanding of the underlying patterns and trends in the data because both sets had similar statistics in many cases. The sample data was carefully selected to ensure that it was representative of the larger population. Its statistical properties indicate that it was a reliable source of information for concluding the entire dataset.

The research discovered the following time-saving potentials:

1. Reduce the number of questions answered within the anxiety checklist to reach the perfect theoretical shortest path - 162,43 seconds;
2. Eliminate the unnecessary questions – 62,82 seconds;
3. Consider the previously answered questions about the same symptoms, thus reducing duplication – 308,13 seconds;
4. Reduce the need for screen transitions – 124,08 seconds; and
5. Eliminate time spent after a confirmed diagnosis – 87,04 seconds.

The discovered potential time-saving aspects sum up to 744,5 seconds. Considering that across all analysed encounters, the mean time duration for an encounter was 1059,43 seconds, the potential time-saving is up to 70,27% of the duration of the encounter.

The time-saving potential could allow psychiatrists to cut back on the data collection process, free up time to focus better on the patient, and start the treatment pathway earlier than the current workflow. However, based on the current study, it is difficult to say that workflow change will automatically reduce the time spent on data capture. Psychiatric practice undoubtedly includes additional factors that affect the time spent answering questions. Enhancements such as leveraging machine learning to improve the questionnaire process and implementing an alternative visual solution for question display and appearance instead of screen transitions may be proper adjustments to improve the data collection process and reduce the overall data collection time.

4.8. Optimizing data capturing and additional findings

The study applies process mining methods to the algorithms and event logs of DocuMental. The approach revealed that there are possibilities for enhancing the data collection process. It seems possible to determine a mental health-related diagnosis more efficiently. The efficiency was demonstrated by comparing the minimum theoretical path to reach a diagnosis against a psychiatrist's path to reach the same diagnosis. As mentioned in chapter 3.8, there is a potential of 11,72% time-saving improvement when considering the lost time on screen transitions, time spent after confirmed diagnosis, and time spent on irrelevant questions. The data, however, does not exactly show how many times the professional has changed tabs in the application, only that the newly answered question was on a different page than the previous one. The number of tab switches may be even more significant; thus, the time spent on screen transitions can be considerably higher.

The data structures and their interconnections were analysed for possible improvements during the research. A previously done research compared rule-based methods to decision tree methods in psychiatric written notes and concluded that rule-based methods are best for information detection tasks. In contrast, decision tree-based methods are preferred for prediction tasks [32]. Another research focused on decision tables and decision trees due to their connection to rule-based systems and concluded that these techniques are essential when building a CDSS [33]. The author of the current research concluded that approaching the data collection problem in psychiatry would best use a decision tree or a rule-based system. No additional research was done on the used algorithms, as DocuMental has already been built on the rule-based system using Java Drools business rules.

The research found 188 questions and their answer options that do not contribute to any of the numerous diagnoses. From that amount, 15 questions within the anxiety checklist do not contribute to any diagnoses. These questions are absent in the data structures of various diagnoses and rule sets, and they merely exist in the checklists as separate questions. The questions likely come from underlying medical guidelines or contribute to the data collection flow somehow. As DocuMental uses so-called opener questions that drive the visibility of other questions, these may be opener questions. However, these algorithms exist in the front-end application of DocuMental [26] and were not analysed in this research. The actual number of unused questions may be lower. As the purpose of these questions is guiding, there exists another discussion point if these should exist at all or if the actual questions contributing to diagnoses should be shown from the start. The author believes these questions could be removed since they do not contribute to the decision-making process. The unused questions amounted to 21,15% of the overall question count. These questions were used on 1286 occasions and have been used an average of 5,67 times per encounter.

It was confirmed by the DocuMental team that, at this time, answers given in the mental status questionnaire carry over to other checklists. However, they must be re-submitted by the professional [26]. The research concluded that on 9232 occasions across all checklists, a question's answer option was chosen, even though a previous question already answered the same question. The answers have been duplicated on 1284 occasions within anxiety-related diagnosis encounters. The arithmetic mean number of answer duplications within the anxiety checklist was 32,92. These actions could have been avoided if previously entered data during the encounter would have been considered. In such cases, the average time that could have been saved by skipping the question or its answer was 308,13 seconds in anxiety-related diagnoses. The mean time between actions, generally, was 9,36 seconds, as recorded in **Table 4**.

In summary, multiple opportunities exist for time-saving improvements in the data collection process in a modern clinical decision support system used in psychiatric care. These opportunities should be addressed in future research and by the DocuMental development team. In addition, further considerations should be discussed if collecting additional data after a confirmed diagnosis is a good practice and whether these data entries will be used in future decision-making processes. It is also possible that these data entries have an output of which this research is unaware.

4.9. Study limitations

This research and the data used for research include a positive bias as the data has been generated mainly by a single professional who is an expert on the platform. Additional psychiatrists have contributed as well. The positive bias within this research primarily stands for a psychiatrist knowing where to click and look for which questions and answers; the same goes for other processes and flows, as one engaged psychiatrist is a creator of the tool, formulas, and rules between data points. Actual data generated by other psychiatrists, who have little to no knowledge of the tool, would have generated data with increasingly substandard statistics. The improvement-related findings on data capturing would have been even more significant.

Numerous steps could have been erroneous during this research, starting from the data collection done by the psychiatrists. Various ethical considerations on the correctness of technicalities have been discussed in chapter 2.3 above.

4.10. Future research

Future research is possible on this subject, as the analysed DocuMental application is still in the prototype phase. Current research can be repeated later with a larger dataset and on other diagnosis groups to reinforce or revise the current research findings. As this application continues developing, processes tend to change, and systems are improved. More research with a larger dataset and improved methods could give an even better insight into the process of using the application.

The author of this research also recommends that future researchers continuing a similar topic have some background knowledge of how psychiatrists work and additional familiarity with FHIR data structures. This is due to the extensive number of technicalities specific to the psychology field overall.

Additional research could be done to analyse how efficient a decision tree-based application would be and compare the results of the newly developed decision tree-based and current rule-based models and their metrics.

As DocuMental currently leverages the rule-based system for its decision-making process, these rules and data structures could be used within machine learning or artificial intelligence frameworks, such as PyTorch, TensorFlow, or the latest GPT models. As process mining and discovery were methods used in this study, they did not include artificial intelligence or machine learning-related opportunities.

Furthermore, additional time studies of psychiatrist encounters and time spent between question answering could be done. This data can then be analysed to see how often and accurately the times recorded by the application are affected by other factors outside the application.

4.11. Final conclusions

The following conclusions based on findings in the research can be drawn for previously stated research questions:

1. Process mining is a valid and suitable methodology to approach problems with ample data in logs and additional data structures.
2. Data quality and inclusion-exclusion criteria are essential to analyse large volumes of data, especially when the researcher needs more confidence in the data quality.

Research question 1: How can decision-making algorithms in a decision support system be mined for the theoretical shortest path to diagnoses?

1. Analysis of the data structures and formulas provides a theoretical shortest path to every diagnosis. The analysis provides a structured set of questions and answers necessary to reach a diagnosis. Visualizing the set offers a good representation of the shortest path.

Research question 2: What is the actual user journey based on the usage logs?

1. The journey of a psychiatrist to reach a diagnosis during an encounter can be analysed from the recorded log of actions. A journey consists of the questions and answers leading to a confirmed diagnosis.
2. The process mining allows for comparing the theoretical path to a diagnosis to the diagnosis pathway chosen by the psychiatrist. The actual user journey and the theoretical shortest paths can be visualized to investigate the improvement potential.

Research question 3: How big is the gap between theoretical minimum and actual usage patterns?

1. An arithmetic mean difference of 14,66 additional questions were answered between the actual and minimal path to an anxiety-related diagnosis.

Research question 4: What are the main contributors to the inefficiency?

1. The main contributors to the gap between the theoretical shortest path to a diagnosis and the actual user journey were unrelated questions, screen transitions, unnecessary questions, and data duplication.
2. Reducing the number of questions answered by a psychiatrist and achieving the minimal theoretical path to a diagnosis could lead to an improvement of 14,66 questions, translating roughly to 162,43-second improvement per encounter only within the anxiety checklist.
3. The mean number of screen transitions during an encounter was 11,28 within anxiety-related encounters. Reducing the number of screen transitions could result in 124,08 seconds saved per encounter.
4. Unused questions were used on average 5,67 times per encounter. The average time spent answering a question was 11,08 seconds, which amounted to 62,82 seconds per encounter.
3. Another possibility is to leverage previously entered answers to other checklists and their questions, which can further improve resources spent on the data collection during an encounter with a psychiatrist. The mean of 32,92 answers (i.e., actions) are repeating across anxiety checklists with a mean of 308,13 seconds potential improvement in the data collection process during anxiety-related diagnoses.

In summary, there existed an opportunity to save time in the data collection process of DocuMental by 70,27%. As the arithmetic mean duration across all encounters was 1059,43 seconds, the potential time-savings could amount to 744,5 seconds. Based on various sources described in the Shortfalls of medical data collection processes chapter, this problem is not unique to the current research and psychiatric field only. Still, it has been noticed by other researchers and professionals as well across different medical fields.

The author believes that selecting an anxiety checklist subset for data analytics and statistics did not affect the outcomes of this research in a significant way. The time-saving opportunities to optimize the data collection process in a clinical decision support system are affected by the same elements of question-answer sequence, screen transitions, unnecessary questions, and data duplication across all checklists. There may be slight differences in the percentage of time-saving opportunities and other statistics, which can be unique to each checklist; however, these do not affect the general behaviour of the user journey within the application.

A study, also mentioned in chapter 1.1, which looked into computerized decision-aid applications, stated that despite the utilization and innovation of IT solutions, modern tools are rarely used in mental health care [20]. Another research described that only a few such applications exist for clinical decision support in psychiatric care, as the technologies are not widely available [21]. This problem persists, as even a unique application like DocuMental is still being used only by a handful of professionals and still looking to get certified [26].

Another research discussed in chapter 1.1 describes that a broad scope of data in psychiatric care exists in a free text and often only investigates a narrow scope of mental health care. These analyses provide valuable insights; however, they often overlook the data collection method and sometimes even go as far as to clean the data only for their research [22]. There is also a need for more evidence to support the usability and usefulness of such systems due to the lack of reliable biomarkers due to a significant portion of the data in free-form text. Due to this, many studies have collected and studied data for their specific purposes, proving the benefits of computerized decision support tools but not contributing to the lack of usable data in mental health care [19]. The current research specifically researched the data collection structures and processes in a modern decision support system and focused on the usability of digital data in mental health care.

The findings of this research are like the problem described by Wachter in the Shortfalls of medical data collection processes. The chapter on improvements done to Mayo Clinic's system allowed the data to be collected on a single screen with 25 clicks instead of the previous 600 clicks and 200 screen transitions and reduced the time taken from 35 minutes to just 5 minutes [16]. That is a reduction of 95,83% in clicks and 85,71% in time taken for the data collection process. Current research concluded that in the DocuMental application, the average number of checklists used per encounter was 2,56, which could be reduced to a single checklist. The actual arithmetic mean of questions answered to reach an anxiety-related diagnosis from the beginning of an encounter was 94; however, this can be opposed by the arithmetic mean of the theoretically shortest path for anxiety-related diagnoses, which was 6,78 – a 92,82% reduction of questions asked by the psychiatrist when compared to the actual mean in perfect conditions. As mentioned previously, the time-saving opportunity is 70,27% of the overall encounter time, which is a good start for the overall reduction in the data collection process.

The research aimed to assess the data gathered with a mental health data collection tool and to demonstrate the possibilities of improving the data collection process in psychiatric decision support. The assessment used process mining techniques to analyse the event logs and discovered improvement opportunities.

5. Summary

This research aimed to evaluate options for improving the data collection process in psychiatric decision support on the example of DocuMental. The author of this thesis utilized process mining techniques to assess the data collection process and demonstrate the potential for improvement. Statistical analysis was done on various metrics computed from the input dataset.

DocuMental, an application used by psychiatrists to conduct patient interviews, was the data source for analysis. The data structures and event logs of DocuMental were researched and analysed. The research analysed the path of questions answered by the psychiatrists and compared the minimum theoretical path to a diagnosis to the actual path taken. The study provided potential insights into bottlenecks and areas of improvement for future research. The research continued focusing on encounters with confirmed anxiety diagnoses assigned by psychiatrists.

The research discovered that potential time-saving aspects include reducing the number of questions, eliminating unnecessary questions, considering previously answered questions, reducing screen transitions, and eliminating time spent after a confirmed diagnosis, which could save up to 70.27% of the duration of the encounter.

The study revealed possibilities for enhancing the data collection and determining a mental health-related diagnosis more efficiently. The data used in the research had a positive bias as it was mainly generated by a single expert psychiatrist who created the tool, formulas, and rules. The research concluded that the data collection process could be optimized significantly and observed multiple possibilities for time-saving improvements in the data collection process of a modern clinical decision support system used in psychiatric care.

References

[1] J. Bulao, “How Much Data is Created Every Day in 2020?,” 10 September 2020. [Online]. Available: https://techjury.net/blog/how-much-data-is-created-every-day.

[2] J. Clement, “Global digital population as of October 2020,” October 2020. [Online]. Available: https://www.statista.com/statistics/617136/digital-population-worldwide/.

[3] B. Marr, “How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read,” 21 May 2018. [Online]. Available: https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read.

[4] J. H. Tibbetts, “The Frontiers of Artificial Intelligence,” *BioScience,* pp. 5-10, 2018.

[5] A. Melnichuk, “How Big Data and AI Work Together,” 21 January 2020. [Online]. Available: https://ncube.com/blog/big-data-and-ai.

[6] E. È. Savander, J. Hintikka, M. Wuolio and A. Peräkylä, “The Patients' Practises Disclosing Subjective Experiences in the Psychiatric Intake Interview,” *Front Psychiatry,* 2021.

[7] M. S. García-Gutiérrez, F. Navarrete, F. Sala, A. Gasparyan, A. Austrich-Olivares and J. Manzanares, “Biomarkers in Psychiatry: Concept, Definition, Types and Relevance to the Clinical Reality,” *Front Psychiatry,* 2020.

[8] P. Taylor, C. A. Pollard and T. M. Flynn, “Use of mobile phones in the behavioral treatment of driving phobias,” *Journal of Behavior Therapy and Experimental Psychiatry,* vol. 23, no. 4, pp. 299-302, December 1992.

[9] R. Bauer, T. Glenn, S. Monteith, P. C. Whybrow and M. Bauer, “Survey of psychiatrist use of digital technology in clinical practice,” *International Journal of Bipolar Disorders,* 2020.

[10] G. Kesari, “AI Can Now Detect Depression From Your Voice, And It’s Twice As Accurate As Human Practitioners,” 24 May 2021. [Online]. Available: https://www.forbes.com/sites/ganeskesari/2021/05/24/ai-can-now-detect-depression-from-just-your-voice/?sh=6735b21f4c8d.

[11] H. Hsin, M. Fromer, B. Peterson, C. Walter, M. Fleck, A. Campbell, P. Varghese and R. Califf, “Transforming Psychiatry into Data-Driven Medicine with Digital Measurement Tools,” *npj Digital Medicine,* vol. 1, p. 37, 22 August 2018.

[12] American Psychiatric Association, “Diagnostic and Statistical Manual of Mental Disorders (DSM–5),” [Online]. Available: https://www.psychiatry.org/psychiatrists/practice/dsm. [Accessed 20 February 2021].

[13] World Health Organisation, “ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines,” 1992. [Online]. Available: https://www.who.int/classifications/icd/en/bluebook.pdf. [Accessed 28 February 2021].

[14] M. Margariti and G. N. Papadimitriou, “Use of informatics technology in psychiatry,” *Psychiatriki,* vol. 23, no. 4, pp. 322-333, October 2012.

[15] E. Chiauzzi and A. Newell, “Mental Health Apps in Psychiatric Treatment: A Patient Perspective on Real World Technology Usage,” *JMIR Mental Health,* vol. 6, no. 4, p. e12292, 22 April 2019.

[16] R. M. Wachter, The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine’s Computer Age, McGraw-Hill, 2015.

[17] A. Männil, “Using the Time and Motion Study Method to Investigate Estonian Family Physicians' Time Utilization During Face-To-Face Patient Visits,” 3 June

2020. [Online]. Available: https://digikogu.taltech.ee/et/Item/3c2b5f25-d25b-4f61-b03a-dee96474702d. [Accessed 13 January 2021].

[18] N. J. Goldenberg, “The Breadth and Burden of Data Collection in Clinical Practice,” *Neurology Clinical Practice,* pp. 81-86, 2016.

[19] T. Aledavood, A. M. T. Hoyos, T. Alakörkkö, K. Kaski, J. Saramäki, E. Isometsä and R. K. Darst, “Data Collection for Mental Health Studies Through Digital Platforms: Requirements and Design of a Prototype,” *JMIR Research Protocols,* vol. 6, no. 6, p. e110, 9 June 2017.

[20] M. Tasma, L. O. Roebroek, E. J. Liemburg, H. Knegtering, P. A. Delespaul, A. Boonstra, M. Swart and S. Castelein, “The development and evaluation of a computerized decision aid for the treatment of psychotic disorders,” *BMC Psychiatry,* vol. 18, 1 June 2018.

[21] A. Suhasini, S. Palanivel and V. Ramalingam, “Multi decision support model for Psychiatry Problem,” *International Journal of Computer Applications,* vol. 1, no. 25, pp. 61-69, February 2010.

[22] L. Moja, K. H. Kwag, T. Lytras, L. Bertizzolo, L. Brandt, V. Pecoraro, G. Rigon, A. Vaona, F. Ruggiero, M. Mangia, A. Iorio, I. Kunnamo and S. Bonovas, “Effectiveness of Computerized Decision Support Systems Linked to Electronic Health Records: A Systematic Review and Meta-Analysis,” *Am J Public Health,* vol. 104, no. 12, pp. e12-e22, December 2014.

[23] L. G. Bergman and U. G. Fors, “Decision support in psychiatry - A comparison between the diagnostic outcomes using a computerized decision support system versus manual diagnosis,” *BMC Medical Informatics and Decision Making,* vol. 8, 8 February 2008.

[24] E. Maron, D. S. Baldwin, R. Balõtšev, C. Fabbri, V. Gaur, D. Hidalgo-Mazzei, S. Hood, M. Juhola, O. Kampman, S. Kasper, H. Kärkkäinen, K. Látalová, M. Lähteenvuo, N. Mastellos and J. McTigue, “Manifesto for an International Digital Mental Health Network,” *Digital Psychiatry,* vol. 2, no. 1, pp. 14-24, 2019.

[25] E. Maron, "Dr Eduard Maron: how DocuMental fills the gap in mental healthcare with innovation," 20 November 2019. [Online]. Available: https://e-estonia.com/documental-healtcare-innovation. [Accessed 15 April 2023].

[26] J. Metsallik, Interviewee, *Chief Technology Officer of DocuMental.* [Interview]. 1 March 2023.

[27] W. van der Aalst, A. Adriansyah, A. K. A. de Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs, A. Burattin, J. Carmona, M. Castellanos, J. Claes, J. Cook, N. Costantini, F. Curbera, E. Damiani, M. de Leoni, P. Delias, B. F. van Dongen, M. Dumas, S. Dustdar, D. Fahland, D. R. Ferreira, W. Gaaloul, F. van Geffen, S. Goel, C. Günther, A. Guzzo, P. Harmon, A. ter Hofstede, J. Hoogland, J. E. Ingvaldsen, K. Kato, R. Kuhn, A. Kumar, M. La Rosa, F. Maggi, D. Malerba, R. S. Mans, A. Manuel, M. McCreesh, P. Mello, J. Mendling, M. Montali, H. R. Motahari-Nezhad, M. zur Muehlen, J. Munoz-Gama, L. Pontieri, J. Ribeiro, A. Rozinat, H. Seguel Pérez, R. Seguel Pérez, M. Sepúlveda, J. Sinur, P. Soffer, M. Song, A. Sperduti, G. Stilo, C. Stoel, K. Swenson, M. Talamo, W. Tan, C. Turner, J. Vanthienen, G. Varvaressos, E. Verbeek, M. Verdonk, R. Vigo, J. Wang, B. Weber, M. Weidlich, T. Weijters, L. Wen, M. Westergaard and M. Wynn, "Process Mining Manifesto," *Business Process Management Workshops,* p. 169–194, 2011.

[28] QPR Software Plc, "The Ultimate Process Mining Guide," [Online]. Available: https://www.qpr.com/process-mining.

[29] The Aphasia Community, "What is Aphasia?," [Online]. Available: https://www.aphasia.com/aphasia-library/what-is-aphasia/. [Accessed 29 April 2023].

[30] B. Dean, "We Analyzed 5.2 Million Desktop and Mobile Pages: Here's What We Learned About Page Speed," 8 October 2019. [Online]. Available: https://backlinko.com/page-speed-stats.

[31] A. Kaoru, G. Naokazu, N. Shin'ya, E. Yoshimichi, T. Tsunehiro and O. Yoshio, "Estimation of the Timing of Human Visual Perception from Magnetoencephalography," *Journal of Neuroscience,* pp. 3981-3991, 2006.

[32] E. Rijcken, K. Zervanou, P. Mosteiro, F. Scheepers, M. Spruit and U. Kaymak, "Machine Learning vs. Rule-Based Methodsfor Document Classification of ElectronicHealth Records within Psychiatry - ASystematic Literature Review," 2 December 2022. [Online]. Available: https://www.researchsquare.com/article/rs-2320804/v1.

[33] B. Silva, F. Hak, T. Guimaraes, M. Manuel and M. F. Santos, "Rule-based System for Effective Clinical Decision Support," *Procedia Computer Science,* vol. 220, p. 880–885, 2023.

Appendix 1 - JSON stream data structure

1. Encounter identificator – 32-character UUID
2. Date and time when the data was read from database
3. Number of entries
4. Entries
   - 4.1. Entry number – consecutively increasing number per entry
   - 4.2. Entry identificator – 32-character UUID
   - 4.3. Date and time of latest change or creation of the entry (*last modified*)
   - 4.4. Entry type
     - 4.4.1. *QuestionnaireResponse* – answer option selection to a question
     - 4.4.2. *Condition* – diagnosis assignment
   - 4.5. Metadata of the checklist connected to the entry
     - 4.5.1. Entry UUID
     - 4.5.2. Naming
     - 4.5.3. Checklist code name, e.g., mental-status-assessment
     - 4.5.4. Display name, e.g., Mental Status Assessment
   - 4.6. Entry status (*in progress*, *completed*, *stopped*, *null*)
   - 4.7. Encounter identifier – same as point 1
   - 4.8. Diagnosis confirmation (only for *Condition* entry type)
   - 4.9. Checklist identificator for diagnosis evidence (only for *Condition* entry type) – 32-character UUID
   - 4.10. Diagnosis status (only for *Condition* entry type) – is the diagnosis *confirmed*, *unconfirmed*, *refuted*?
   - 4.11. Diagnosis internal code (only for *Condition* entry type)
   - 4.12. Entry (only for *QuestionnaireResponse* entry type)
     - 4.12.1. List of questions within the checklist and their descriptions e.g., `{ „linkId“: „123“, „text“: „Eye contact“ }`
     - 4.12.2. Answer options selected for the question – a new element “answer” is added to the data structure in previous point. This contains the answer option number.

Appendix 2 - Checklist usage across encounters