

Using process mining to detect workarounds that are used to reach an alternative goal

Victor A. van Andel

A thesis presented for the degree of
Master of Science (MSc)



Utrecht University

Utrecht University
Department of Information and Computing Sciences
Master's programme: Business Informatics
The Netherlands
10th August 2020

First supervisor: Iris Beerepoot
Second supervisor: Inge van de Weerd
Third supervisor: Xixi Lu

Using process mining to detect workarounds that are used to reach an alternative goal

Victor A. van Andel

Abstract

Modern day healthcare revolves around Health Information Systems (HISs) and Electronic Health Records (EHRs) that are stored in the cloud. Accountability has become a hot topic as patients want to know who accessed their record and why. Anecdotal evidence points to a risk where the justification for accessing records is not always validated. For this phenomenon I coin the term Alternative Goal Workaround (AGW). The research goal of this design science study is to explore this concept and to narrow the research gap in quantitative workaround detection. In a case study at a Dutch hospital, five AGWs are identified through stakeholder interviews. The PM²4AGW methodology presents a way to distinguish between workarounds and legitimate process instances through the principle of exclusion. In iterative analysis cycles, patterns of legitimate behavior are captured to reduce false positives and detect workarounds with increasingly high precision, as demonstrated on the five workarounds at the case study hospital. By capturing knowledge about the AGWs in a snapshot, non-malicious AGWs can be addressed to improve the process, after which malicious AGWs will stand out even more. The PM²4AGW methodology has empowered the hospital to more efficiently catch and sanction unlawful use of EHRs, making the hospital a safer place for patients.

Keywords

Workarounds, alternative goal, business processes, process mining, methodology, security, privacy, transparency, healthcare, health information systems, electronic health records, snapshot, awareness.

Contents

1	Introduction	3
1.1	Problem Statement	3
1.2	Research Questions	4
1.3	Contributions	5
1.4	Thesis outline	6
2	Background Literature	7
2.1	Workaround Definitions and Types	7
2.2	Views on Workarounds	9
2.3	Concepts related to Workarounds	11
2.4	The use of Process Mining	12
2.5	Patient Privacy	14
3	Research Approach	17
3.1	Design Science	17
3.2	Case Study: Rationale & Design	17
3.3	Case Study: Context	18
3.4	Data Collection & Analysis	18
3.4.1	Interviews	19
3.4.2	Event data	19
3.4.3	Focus group	20
3.5	Assessment of Threats to Validity	20
4	Results	22
4.1	Alternative Goal Workarounds	22
4.2	The PM ² 4AGW methodology	24
4.3	Measuring Workaround Detection Performance	27
4.4	Snapshot for AGWs	31
5	Discussion	33
6	Conclusion	35
	Acknowledgements	35
	References	36
A	Information security practices	40
B	Contact moments with stakeholders	42
C	Classic workarounds	43

1 Introduction

The emergence of Health Information Systems (HISs) has enabled the digital availability of information that is needed by health workers to accomplish their job. Information about a patient is found in the Electronic Health Record (EHR), which has become an integral part of modern healthcare (Borycki et al., 2011). With the arrival of EHRs, elaborate security and privacy requirements must be formulated to protect the sensitive data that they contain (Rodrigues et al., 2013). Sometimes, it can be costly or even impossible for employees to comply with all security policies and mechanisms while carrying out their job effectively (Kirlappos et al., 2015). Meanwhile, it is challenging to design systems to support healthcare processes as they are highly complex and dynamic (Mans et al., 2008; Rebuge & Ferreira, 2012) and health workers can take on multiple roles at different times (Koppel et al., 2015). It is hard to take everything into consideration during the design phase of a system, and changes often need to be made as soon as users get their hands on a new system (Arduin & Vieru, 2017). Due to these factors, there is often some misalignment between HISs and the processes they aim to support (Beerepoot, Ouali, et al., 2019; Beerepoot & van de Weerd, 2018; Koppel et al., 2015; Arduin & Vieru, 2017). This misalignment can lead health workers to deviate from the processes as intended in an effort to accomplish their work despite the limitations of the HIS. This phenomenon is called a workaround, and workarounds are a common occurrence in the healthcare sector (Kobayashi et al., 2005). Existing definitions of workarounds focus on how people deviate from a designed path to reach the same goal, but it could also be the other way around. What if people exploit a designed path to reach an alternative goal?

1.1 Problem Statement

Anecdotal evidence suggests that workarounds where people exploit a designed path to reach an alternative goal pose a significant risk in healthcare. Health workers can access a large range of patient records through paths that are designed for specific situations. This is a vulnerability, as it is labor-intensive to investigate whether these designed paths were taken for rightful use or not. Misuse of these designed paths to reach malicious or non-malicious alternative goals can currently fly under the radar as it is infeasible to manually monitor all access to patient records.

Thus far, the key method for detecting workarounds was through qualitative research techniques such as observing and interviewing staff in their work environment (Beerepoot, Ouali, et al., 2019; Koppel et al., 2008; Koppel et al., 2015; Cresswell et al., 2017). However, the qualitative approach is labor-intensive and relies on the willingness of participants to expose their work practices to an observer (Beerepoot et al., 2018). Qualitative research might also be influenced by the researcher’s perception or be inconsistent when multiple researchers partake. It can also be challenging to determine when qualitative data collection is (sufficiently) completed.

As an alternative, it has been suggested as future research direction to detect workarounds using a quantitative approach, more specifically, by applying process mining (Beerepoot, Koorn, et al., 2019; Beerepoot & van de Weerd, 2018). Process mining is a family of techniques that bridges the gap between traditional model-based process analysis and data-centric analysis techniques (van der Aalst, 2016). Beerepoot & van de Weerd (2018) mention that it may help with automatically detecting and monitoring workarounds. Beerepoot, Koorn, et al. (2019) propose that it may supplement qualitative methods, though they note that many types of workarounds cannot yet be detected using data analysis techniques. This is supported by an early attempt by Outmazgin & Soffer (2013), which showed that not all generic types of workarounds can be identified using process mining without the use of additional domain knowledge. On a similar note, Koppel et al. (2015) wrote that it requires more than just an analysis of computer rules and access logs with permission levels to understand the circumventions of cybersecurity in a healthcare setting. The addition of qualitative research helps to understand the creativity, flexibility, and motivation of well-intended people in accomplishing their work despite obstacles posed by security technologies and measures (Koppel et al., 2015). These remarks show that detecting workarounds quantitatively should not be underestimated and requires a good understanding of the complex problem domain that is healthcare.

Quantitative workaround detection is a gap in workarounds research as it remains largely unexplored, with the study by Outmazgin & Soffer (2013) being a rare exception. At this point in time, there is still insufficient knowledge to determine how effective quantitative workaround detection is and what the best way is to develop and optimize it. It is fair to note that quantitative detection might not be well suited for all kinds of workarounds. For example, some workarounds might circumvent the system altogether and leave no trace in the event logs. Nevertheless, there could be much to gain in working towards a quantitative approach. Once a reliable methodology has been discovered for quantitative detection of certain workarounds, it can be re-used in similar environments with some slight changes. It can also be employed on a larger scale to automatically monitor the workarounds without requiring many additional resources, and it is transparent, objective, and evidence based.

1.2 Research Questions

The problem statement announced the gap that exists in workarounds research. Where qualitative detection is limited by human researchers, quantitative detection of workarounds could offer many benefits such as scalability and transparency. The need for hospitals to account for the access to sensitive patient information further underlines the practical usefulness for a reliable methodology to detect and monitor workarounds that are used to reach an alternative goal. To address the gap, I have formulated the following main research question for my thesis:

How can process mining be used to detect workarounds that are used to reach an alternative goal?

To answer all aspects of this question, it is broken down into the following set of sub-questions that will first be answered individually. After presenting each sub-question, I will give a brief summary of my approach for answering the question.

SQ1: Which and what kinds of alternative goal workarounds are there?

In order to gain a better understanding of alternative goal workarounds, I will conduct interviews with the people who are involved in them. This includes those who practice them and those responsible for the proper execution of the process. Then I will create a classification of the workarounds that I identify through these interviews.

SQ2: What is a suitable methodology for detecting the workarounds?

I will investigate the effectiveness of quantitative workaround detection by looking for a suitable methodology to detect the identified workarounds using process mining. The methodology needs to capture a way to extract the right data and process it into an event log from which workarounds can be distinguished from legitimate process instances.

SQ3: How can we measure the performance of the workaround detection?

After establishing a way to detect the workarounds, I will shift my attention to measuring the performance, covering how to handle false positives and false negatives and how to know when the detection has been optimized. Additionally, I will provide an overview of the performance of the specific workarounds that I studied.

SQ4: What information should be captured about the workarounds to initiate process improvement?

Discovering workarounds and being able to detect them systematically may be an incentive for improving the processes in which the workarounds occur. I will organize a focus group with experts to discuss what information about the workarounds should be captured in order to initiate process improvement.

1.3 Contributions

In this thesis I study a new phenomenon, the Alternative Goal Workaround (AGW). I provide a definition for it and identify five examples of AGWs that are practiced daily at a Dutch hospital. My main contribution lies in addressing the research

gap of using quantitative methods to detect workarounds systematically. Through design science, I have developed the PM²4AGW methodology for detecting AGWs with process mining, which was demonstrated on the five identified workarounds. For evaluation, I have devised a way to measure the performance of the workaround detection which shows how the methodology leads to an improvement with each iterative cycle. By defining what information needs to be captured about AGWs, I took another step forward in formalizing them and using them to initiate process improvement.

Not only are these theoretical contributions, they are also suited for practical use as demonstrated in the case study. For the five studied workarounds I was able to create an accurate detection that will be implemented and used for operational support. It empowers the privacy officer of the hospital to more efficiently catch unlawful use of EHRs. The workaround detection mechanism eliminates many false positives by automatically checking for known events that legitimize the process. This reduces the amount of cases that the privacy officer needs to investigate manually and increases the probability of catching malicious activity.

1.4 Thesis outline

The next chapter contains a review of literature that provides a background for the topics discussed in the thesis. The third chapter will cover the research approach, in which I will elaborate on the way that the research was conducted and the decisions behind it. In chapter four, I will present the results for each of the sub-questions. The discussion in the fifth chapter reflects on the research and its limitations, positions the study, and provides some suggestions and recommendations for future research on this topic. The sixth and final chapter contains the conclusion, which answers the main research question and wraps up the thesis.

2 Background Literature

Relevant background literature was gathered in several ways. The snowball method was applied using a few key documents, namely: Alter (2014), Beerepoot, Ouali, et al. (2019), and Beerepoot & van de Weerd (2018). Additionally, the following search terms were used on Google Scholar: *workaround*, *workaround process mining*, *process mining*, *workaround security*, *workaround privacy*, and *patient privacy*. A selection of relevant literature was made from these sources based on the criteria that the source was written in the English language, digitally accessible, and stated something of relevance to at least one of the literature topics. Moreover, recent works were preferred over older ones. Studying the literature helped to provide a solid theoretical background for carrying out this study, which is presented in this chapter.

2.1 Workaround Definitions and Types

In his "Theory of workarounds", Alter (2014) brought nearly all previous research on workarounds together in an extensive literature review, consisting of more than 300 articles. Analyzing these articles yielded not only different viewpoints and perspectives on workarounds, but also a large diversity of examples of them. During this analysis, he established that there was not yet a comprehensive theory of workarounds that was broad enough to cover most of the known examples and the situations in which they occur. To fill this gap, Alter (2014) proposed his own theory, built on the theory of planned behavior, bricolage and improvisation, agency theory, and work system theory. Alter defines a workaround (in an organizational setting) as follows:

"A workaround is a goal-driven adaptation, improvisation, or other change to one or more aspects of an existing work system in order to overcome, bypass, or minimize the impact of obstacles, exceptions, anomalies, mishaps, established practices, management expectations, or structural constraints that are perceived as preventing that work system or its participants from achieving a desired level of efficiency, effectiveness, or other organizational or personal goals" (Alter, 2014, p. 1044).

This definition is more inclusive than most of the preceding definitions and has since been adopted in many other works on workarounds (Beerepoot, Koorn, et al., 2019; Beerepoot, Ouali, et al., 2019; Burns et al., 2015).

As a result of their literature review, Ejneffjäll & Ågerfalk (2019) created an overview of the five most common definitions of the term 'workaround', including a visual representation which is shown in Figure 1. These definitions can also be thought of as different types of workarounds with certain properties that may or may not be present such as the workaround being intentional, goal-oriented, or whether they overcome a workflow block. Definition 5 represents the definition by Alter

(2014) which is intentional, goal-oriented and may or may not exist in order to overcome a workflow block.

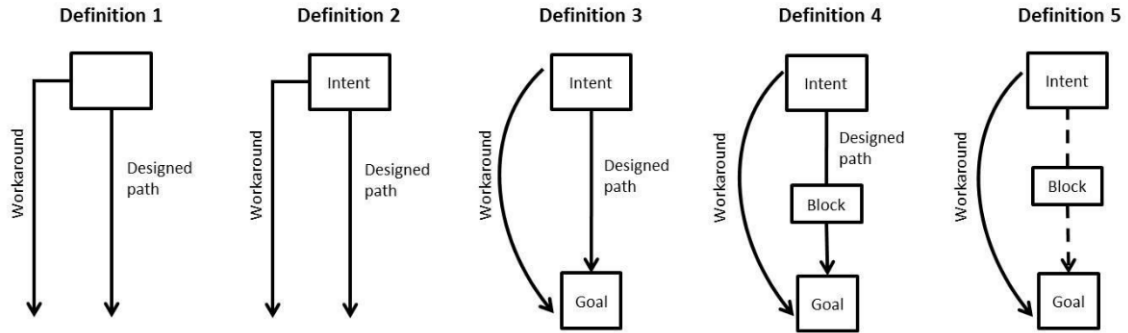


Figure 1: Visualization of the five definitions of the term 'Workaround' (Ejnefjäll & Ågerfalk, 2019)

In all of these definitions, the actor deviates from the designed path and in three of them, it is explicitly stated that the actor does so in order to achieve the same goal in a different way. However, anecdotal evidence from the healthcare sector suggests that it could also be the other way around. An actor may exploit a designed path to reach an alternative goal. Therefore, I coin the term 'Alternative Goal Workaround' (AGW) to define a new type of workaround that I will explore in this thesis. With this workaround, the user intentionally uses (a part of) the designed path to reach an alternative goal (either malicious or non-malicious). The visualization in Figure 2 shows both the 'classic' workaround and the alternative goal workaround.

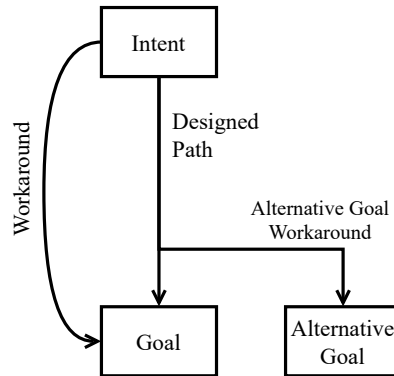


Figure 2: Visualization of the Alternative Goal Workaround

The visualization shows the designed path that leads from the intent to the goal, depicting the intended process. The curved arrow on the left shows the 'classic' workaround (by Definition 3 in Figure 1) which circumvents the designed path to reach the same goal. The remaining arrow starts out by following the designed path,

but then branches out to the right in order to reach the alternative goal, depicting the newly defined alternative goal workaround.

2.2 Views on Workarounds

Like his definition, the theory of workarounds by Alter (2014) was created to encompass the large variety of phenomena that are called workarounds. It is a 'process theory' that describes how a workaround is developed, the factors that contribute to it, whether it might be appropriate, and its consequences. The usefulness of the theory for structuring and understanding workarounds has been endorsed and it has served as a foundation to further the knowledge of workarounds (Arduin & Vieru, 2017; Röder et al., 2014).

Alter (2014) summarizes his literature review in a 'five voices model' that characterizes the *phenomena associated with workarounds*, *types of workarounds*, *direct effects of workarounds*, *perspectives on workarounds*, and *organizational challenges and dilemmas related to workarounds*. Alter (2014) concluded that attitudes towards workarounds vastly differ and that positive, neutral, and negative views all appear in the literature. These are portrayed in the *perspectives on workarounds* dimension of the five voices model. From a positive perspective, workarounds can be viewed as creative acts that allow employees to fulfill their responsibilities despite the obstacles that get in their way. A negative perspective is that workarounds are viewed as a form of resistance to authority and newly introduced technologies.

A main risk of workarounds is that they often introduce security vulnerabilities as a side effect (Arduin & Vieru, 2017). More specifically, unattended workarounds can be harmful to the overall organizational privacy compliance (Murphy et al., 2014). Workarounds introduce informal workflows which makes it almost impossible to guarantee information security and patient privacy (Burns et al., 2015).

The upside is that when an organization becomes aware of workarounds and the weaknesses in their work system, it is also an opportunity for them to mitigate the associated risks and improve the work process (Arduin & Vieru, 2017; Cresswell et al., 2017). Another possible advantage of a workaround is the expected efficiency gain (Röder et al., 2014). According to the model by Röder et al. (2014), this expected efficiency gain increases the management's willingness to tolerate the workaround. On the other hand, the exposure to compliance risk and perceived process weakness decrease the willingness to tolerate the workaround.

There have also been accounts of when workarounds threatened the safety and well-being of patients, especially in the context of medication prescription systems. In one case, prescribers got used to ignoring the majority of safety alerts from the system (Vogelsmeier et al., 2008). In another case, clinicians were forced by the system to order a second dose of blood thinners for a patient (lethal if the patient actually received it) and re-log in the system to cancel the order immediately afterwards (Koppel et al., 2015).

One might wonder how and why workarounds come into existence. According to Alter (2014), workarounds are caused by two general problems that may occur individually or in combination: *obstacles to doing work in a preferred manner* and *misalignment of goals and incentives of actors, principals, and other stakeholders*. Similarly, a common notion is that workarounds emerge when there is a misalignment between an information system and the business process that it must support (Beerepoot, Ouali, et al., 2019; Beerepoot & van de Weerd, 2018; Koppel et al., 2015). Often, this misalignment originates from the design or implementation phase of the system (Vogelsmeier et al., 2008; Outmazgin, 2013; Arduin & Vieru, 2017). The EHR is only one representation of reality and might differ from the patient’s physical reality or the clinician’s mental model of the patient’s condition or not be able to represent them accurately at all (Smith & Koppel, 2014). The emergence of workarounds seems hard to avoid, as not everything can be taken into consideration during the design phase of the system (Arduin & Vieru, 2017). Sometimes, workarounds may emerge when there are policies in place that are impossible to comply with and get work done at the same time (Kirlappos et al., 2015). Besides that, the other main reasons why employees could resort to workarounds according to Kirlappos et al. (2015) are a *lack of awareness* and *high compliance costs*. Andrade et al. (2016) also propose five factors that trigger noncompliant behavior and categorize them into intended and unintended noncompliance. Unintended factors are a *lack of knowledge* and *carelessness*, whereas intended factors are *desire to improve process outcome*, *desire to prevent future mishaps*, and *desire to avoid tedious tasks*. They concluded that intended noncompliance, to which they count workarounds, generally has a more positive effect on the business process outcomes than unintended noncompliance. An exception to this was the factor *desire to avoid tedious tasks*, which had a predominantly negative effect on the business process outcomes (Andrade et al., 2016).

To summarize, workarounds are goal-driven deviations from work protocol that employees take in an attempt to do their job despite the inadequacy of their work system. Workarounds emerge when such a work system is misaligned from the business process and forms an obstacle for the employee to achieve his goals. Workarounds can have positive effects on the organization, such as gains in efficiency, and they may sometimes even be necessary in order for the process to be executed at all. Meanwhile, they can have some severe negative effects as well, especially with regard to compliance with company policy, security, and even the safety and well-being of patients. It is worth investigating workarounds in organizations, as it can allow them to capitalize on the positive effects or tackle the negative effects, or both. Workarounds provide useful information that can lead to both process improvement and better process management.

2.3 Concepts related to Workarounds

First and foremost, workarounds are related to the research area of compliance management. Alter (2014) incorporated *compliance or noncompliance with management intentions* as one of the direct effects of workarounds in his five voices model. Outmazgin & Soffer (2013) focused on two of the main activities that are seen in compliance management, which are compliance checking and compliance improvement. As the names indicate, compliance checking is concerned with determining whether a process meets certain constraints, and compliance improvement with modifying the process to improve the compliance (Outmazgin & Soffer, 2013). When compliance checking results in the detection of noncompliant behavior, this often leads to compliance improvement as a countermeasure. Compliance checking can be done both forward and backward, where forward compliance checking assesses the design and implementation of processes and backward compliance checking purely focuses on detecting noncompliance in running processes (Outmazgin & Soffer, 2013; Ramezani et al., 2012). Workarounds are considered by Outmazgin & Soffer (2013) and Andrade et al. (2016) to be a specific form of noncompliance, which is a perspective that differs from Alter's. Furthermore, Andrade et al. (2016) conclude that workarounds as per the definition of Alter, are an intentional form of noncompliance, as they are goal-driven. According to Arduin & Vieru (2017), intentional workarounds can either be malicious or non-malicious. So altogether, workarounds can be viewed as a form of intentional noncompliance which may or may not be malicious.

Besides compliance checking, there is also conformance checking. Compliance checking focuses on whether a process complies with specific constraints. Conformance checking on the other hand, uses an initially defined process model and detects deviations from that in practice (van der Aalst et al., 2011). However, the two terms seem to often be used interchangeably.

Kirlappos et al. (2014) identified a new type of behavior that is in between compliance and non-compliance with an organization's security policy and called it 'shadow security'. They defined it as *"instances where security-conscious employees who think they cannot comply with the prescribed security policy create a more fitting alternative to the policies and mechanisms created by the organization's official security staff"* (Kirlappos et al., 2014). In practice, the occurrence of these 'security workarounds' is a compromise between employees accomplishing their job and managing the risks that come with it (Kirlappos et al., 2015). In other words, this type of workaround actively attempts to mitigate the security risks that it causes.

Vogelsmeier et al. (2008) introduce the concept of 'work flow blocks', which are interruptions to a work flow that are caused by technology. A common type of work flow block is when an HIS gives an alert that asks the health worker to reconsider an action that is potentially unsafe, though they can also be caused by various other functionalities of the system. Despite their positive intentions, work flow blocks are often considered a distraction that causes errors instead of preventing them. To avoid them as much as possible, health workers often engage in workarounds as a

form of first-order problem solving (Vogelsmeier et al., 2008).

Cresswell et al. (2017) distinguish between formal and informal workarounds. Informal workarounds are not approved by management (or even unknown to them) and solve shortcomings in the design of a work system. Formal workarounds on the other hand are accepted and sometimes even promoted by management as they are necessary for the organization to function. Beerepoot, Ouali, et al. (2019) argue that when a workaround is used repetitively and becomes established practice, it reaches the end of its lifecycle and should no longer be considered a workaround.

Some of the concepts discussed in this section, among others, were also covered by Röder et al. (2016) in their classification in response to the lack of a general consensus on workarounds. They propose that workarounds types mainly need to be differentiated based on their intention and outcome.

2.4 The use of Process Mining

As an alternative for qualitative workaround detection, it has been suggested to detect workarounds using a quantitative approach by applying process mining (Beerepoot, Koorn, et al., 2019; Beerepoot & van de Weerd, 2018). Process mining is a family of techniques that bridges the gap between traditional model-based process analysis and data-centric analysis techniques (van der Aalst, 2016).

The only inescapable prerequisite for applying process mining is that the system produces event logs (Mans et al., 2008). With a case study at an academic hospital, Mans et al. (2008) demonstrated the applicability of process mining in the healthcare domain. Though the complexity of the domain and the lack of structure in the processes made it challenging, they were able to derive understandable models from the event logs that gave insight into the processes. Along with 73 other papers, the study by Mans et al. was included in a literature review by Rojas et al. (2016). The review covered all well-reported case studies of process mining in healthcare and analyzed several of their characteristics. These included the processes and data types being analyzed, the tools and techniques used, and the analysis strategies. It provides a good overview of the different approaches to process mining in healthcare. Mans et al. (2015) describe what kind of processes and data one may encounter in healthcare. Their healthcare reference model aims to support data extraction for process mining.

Rebuge & Ferreira (2012) underlined both the importance and the difficulty of analysing healthcare processes. Because traditional Business Process Analysis (BPA) techniques were lacking, they looked at process mining as a promising alternative to perform BPA with. They proposed a methodology based on the general methodology for performing process diagnostics with process mining by Bozkaya et al. (2009). After preparing and inspecting the event logs, the adaptation of Rebuge & Ferreira (2012) adds the sub-process of performing sequence clustering before moving on to the control flow, performance, or organizational analysis. The clustering

technique helps to cope with the large amount of noise that is often found in health-care data. It is also key to distinguishing regular behavior from process variants and infrequent behavior. They demonstrated the effectiveness of their methodology through a tool that they implemented during a case study at a hospital emergency service.

Mans et al. (2008) divided the domain of process mining into three main types that are defined by the nature of their activities. Originally, process mining focused on *discovery*, which requires no a-priori model of the process. Its goal is to construct a model that describes the behavior that is observed in the event logs in order to extract knowledge from the process. *Conformance* on the other hand, uses an a-priori model that describes how the process should be executed. The event logs are checked against the model to detect deviations from the prescribed process – in other words, conformance checking. Lastly, *extension* also uses an a-priori model but is only concerned with enriching it with the data from the event logs. The three types of process mining by Mans et al. (2008) are visualized in Figure 3.

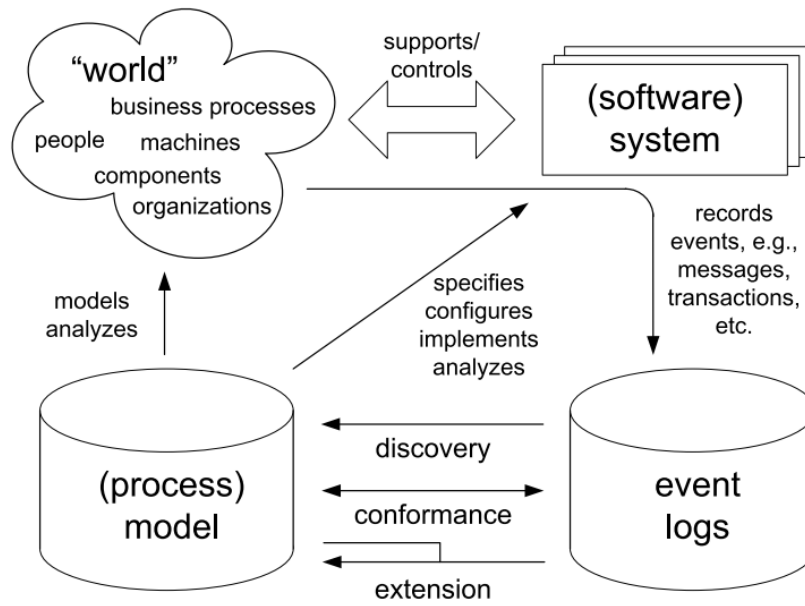


Figure 3: Three types of process mining (Mans et al., 2008)

In their case study, Mans et al. (2008) focus on the discovery part of process mining. After pre-processing the logs for simplification, they are analyzed from the three major perspectives in process mining. The most common one is the *control flow* perspective, which automatically derives a process model from the event logs by applying an algorithm. The *organizational* perspective focuses on the relations between actors to sketch a social network in order to discover patterns of collaboration. The *performance* perspective looks at the duration of different parts of the

process which can serve to identify bottlenecks in the process.

Outmazgin & Soffer (2013) were one of, if not the first to apply process mining to the identification of workarounds specifically. They view business process workarounds as a specific form of incompliance that is intentional. For example, when an employee is aware of internal policies and external regulations, but decides to work around them nonetheless. Additionally, they view quantitative identification of workarounds as a specific form of backward compliance checking. Their study utilized a list of six known generic types of workarounds, that were previously identified in a qualitative exploratory study (Outmazgin, 2013). After discussing the workaround types, Outmazgin & Soffer (2013) theorized how each of them might be detected using process mining through the patterns they should emit in the event logs. These log patterns come in the form of control flow compliance rules defined by Ramezani et al. (2012). They used rules from the *existence*, *precedence*, *response*, and *between* categories and implemented them as filters over the event logs, where every process instance that contained a pattern at least once was considered a workaround of that type. Through this method, they were able to detect four out of the six workaround types, which they successfully demonstrated in multiple real cases. They report the percentage of process instances in which each of the four workaround types occur. However, they do admit that their findings could be prone to both false positives and false negatives. The remaining two workaround types leave no recognizable trace in the event logs and could therefore not be identified using process mining on a generic level. Additional domain knowledge might help identify the remaining two types through more (case-)specific patterns. This goes to show how quantitative techniques can be limited in their ability to identify workarounds, especially as some workarounds might evade the logging altogether.

2.5 Patient Privacy

As mentioned before, workarounds are known to introduce security vulnerabilities in organizations (Arduin & Vieru, 2017). When left unattended, they can be harmful to the organizational privacy compliance (Murphy et al., 2014). The organization’s information security and patient privacy cannot be guaranteed as a result of informal workflows introduced by workarounds (Burns et al., 2015).

After conducting a literature review, Leino-Kilpi et al. (2001) concluded that “the concept of privacy is highly complex and involves different perspectives and dimensions, and there is no single universal definition of privacy”. Out of the four dimensions of privacy defined by Burgoon (1982), the *informational* privacy dimension is the most relevant to the use of EHRs in healthcare. The informational dimension is concerned with the rights of a person regarding the collection and processing of information about them. This aspect becomes more and more prevalent with the continuing growth of information processing such as seen especially in healthcare (Leino-Kilpi et al., 2001). The standard for classification of purposes

for processing personal health information in health informatics (International Organization for Standardization, 2011) used the definition for privacy that originally appeared in the standard for vocabulary in information technology (International Organization for Standardization, 1998). This definition of privacy is as follows: "freedom from intrusion into the private life or affairs of an individual when that intrusion results from undue or illegal gathering and use of data about that individual". As individuals can have a different interpretation of what counts as intrusion, privacy is subjective up to a certain degree. Legislation ever evolves as it attempts to determine how the capturing and sharing of data should be handled.

There is a lot of legislation for information security that is applicable to Dutch hospitals. First of all, there is the European General Data Protection Regulation (GDPR) which is concerned with the processing of personal data of all citizens of the European Union (EU). The Dutch Data Protection Authority (Autoriteit Persoonsgegevens) is the independent body that has been appointed by the government to supervise the compliance with this regulation. Additionally, healthcare organizations in the Netherlands need to comply with the NEN 7510 standard developed by the Netherlands Standardisation Institute. This standard is an adaptation of the Dutch Code for Information Security that has been made specifically for the healthcare sector to improve the clarity and lower the threshold for compliance. At the time of writing, the NEN 7510 is supplemented by the NEN 7512, the NEN 7513, and the NEN 7521. Within the context of this thesis, the NEN 7513 is especially relevant as it prescribes health organizations to collect logs of all actions that are performed with respect to EHRs.

It is also important to take into account that different rules might hold for different parts of the EHR. Murphy et al. (2014) distinguish between three kinds of patient information: protected health information (anything about a patient's medical status), personally identifiable information (anything that could identify the patient), and generalized information (anything else). A different level of discretion may be required depending on the kind of information that is dealt with. This could be embodied by restricting certain parts of the records for certain kinds of employees.

The most common form of system access control in hospitals is user authentication (Murphy et al., 2014). By requiring users to be logged in, every action that is performed within the system can be traced back to a user. So even when a malicious action is not prevented, it can be discovered, and the offender can be held accountable. Simply knowing that every action can be traced back to you might prevent people from acting inappropriately (e.g. snooping around in their neighbor's record). However, problems may arise when employees share passwords with each other or leave their workstation unattended without locking their device. These are also known examples of workarounds in the authentication and de-authentication processes respectively (Koppel et al., 2015). In order to guarantee accountability, it is essential to educate employees on proper authentication practices.

Another critical issue that arises with access control is determining which permissions should be given to which user – in other words, user provisioning. Give a user too many permissions and questions will be asked. Give him too little and he will not be able to do his job. User provisioning is quite complicated in reality, especially in healthcare where information needs often shift (Koppel et al., 2015). Health workers can have multiple roles, and even move between different hospitals that use the same system. Patients can also move from one department to the other or be treated at multiple ones simultaneously. In some situations, it can be lifesaving to quickly gain access to a patient’s data. Therefore, some hospitals employ a ‘break the glass’ mechanism, which health workers can use to transcend their permissions and gain access to a specific record in case of an emergency (Lovis et al., 2007). At the hospital studied by Lovis et al. (2007), the use of this mechanism was meant to be kept to a minimum and be reserved for real emergencies. To achieve this, the mechanism prompted the user with a text field to justify its use which was routinely investigated at the end of the month. Using the ‘break the glass’ mechanism in situations that it is not meant for can be seen as a workaround. When it is used as a workaround too often, investigating its use becomes more and more burdensome.

3 Research Approach

This chapter will go into the design of the research and the decisions that were made to arrive there. It will first cover the research methodology that was used. Then, some details will be provided on the design and the context of the case study. Following that, I will discuss how data was collected and analyzed during the case study. Finally, an assessment is provided of the threats to the validity of the research.

3.1 Design Science

I structured my research using the Design Science Research Methodology by Peffers et al. (2007). This methodology is well-suited as it provides a framework of steps for designing an artifact and validating it by demonstrating its effectiveness when solving a problem in its context. Figure 4 illustrates how the steps were given shape for this research project.

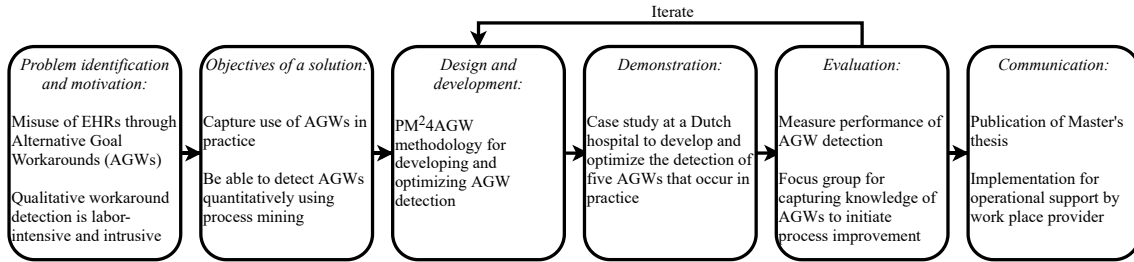


Figure 4: Design Science Research Methodology (after Peffers et al. (2007))

The identified problem is the misuse of EHRs through AGWs and the shortcomings of qualitative workaround detection. The objective of the study is to identify several AGWs that are used in practice and detect them quantitatively using process mining. This is achieved by developing a methodology and demonstrating and evaluating its effectiveness on the identified workarounds during an extensive case study.

3.2 Case Study: Rationale & Design

I opted to conduct a case study because it is a flexible research approach which is able to cope with the complex and dynamic characteristics of real world phenomena (Wohlin et al., 2012), such as those found in healthcare (Rebuge & Ferreira, 2012). This flexibility allows for the integration of different research methods such as interviews and quantitative data analysis. Instead of drawing a sample from some population and carrying out a controlled experiment, the phenomena are studied in their context from various perspectives. By introducing the designed artifact in that context, its impact can be demonstrated and evaluated with stakeholders.

The case study uses a single-case design (Yin, 2014). The main rationale for this being that it is a common and representative case (Yin, 2014). The main HIS used at the case study hospital, Chipsoft HiX, is the market leader in the Netherlands. Besides that, the topic of patient privacy is currently relevant at the hospital as they have recently started monitoring the access to patient records more closely. Combined with the relationship of trust they have with Intermax, this formed a good opportunity to conduct this case study. Furthermore, it is an embedded case study, since it integrates both qualitative and quantitative methods to investigate multiple units of analysis, namely the workarounds (Yin, 2014; Scholz & Tietje, 2002).

3.3 Case Study: Context

The case study was conducted at a large general hospital situated in the Netherlands. The hospital is a client of Intermax Cloudsourcing BV, which was my workplace provider during this thesis project. Intermax is specialized in hosting Chipsoft HiX, which includes the EHRs for all the patients, in a secure private cloud environment. Intermax recently launched an EHR audit module called EPD Insights in collaboration with their sister company NFIR.

For the purpose of this study, I was given access to all audit logging of the EHRs under the conditions of an Non-Disclosure Agreement (NDA). The hospital also gave me access to documentation on their privacy policy and processes and allowed me to conduct interviews with several stakeholders who were of interest for the case study. For additional context of the case study hospital, Appendix A includes an overview of their information security policy, following from an interview and documentation.

The biggest stakeholder at the hospital was the Privacy Officer (PO), whose job the workaround detection will mainly support. She played an important role in both identifying workarounds at the start and optimizing the detection of those workarounds. Other involved stakeholders included the Chief Information Security Officer (CISO), the Chief Medical Information Officer (CMIO), the Chief Nursing Information Officer (CNIO), a system administrator, a physician, and a nurse at an outpatient clinic.

3.4 Data Collection & Analysis

This section covers the collection and analysis of data throughout the case study. Shortly after starting the case study, the COVID-19 outbreak in the Netherlands began, which prevented me from visiting the hospital in person from that point onward. Therefore, some changes needed to be made in the methods for collecting qualitative data. The section is divided into the three methods through which data was collected: interviews, event data extracted from the HIS, and the focus group. Appendix B provides an extensive overview of my various contact moments with stakeholders at the hospital.

3.4.1 Interviews

At the beginning of the case study, I interviewed the CISO and the Privacy Officer on how the privacy of patients is guaranteed and the various information security policies and practices at the hospital. To provide some more insight into the information security landscape of the hospital, the results of this interview can be found in Appendix A.

Due to the COVID-19 outbreak in the Netherlands it was not possible to conduct the interviews for workaround identification in person at the case study hospital. Instead, I created an interview protocol to adapt to the situation. After receiving permission for the interview from the participant, I would e-mail them instructions for the interview. I sent them a single page document that explained the purpose of my study, the concept of workarounds, and what information I would like to gather from the interview with them. I kindly asked them to study the document and think about my questions for a couple of days before scheduling a call. This gave them some time to recognize workarounds in their work environment and come up with examples. Then, during a scheduled call, they would relay their findings to me and I could ask follow-up questions to get a better understanding of the potential workarounds that they encountered. I focused on finding out whether and how I might be able to distinguish between workarounds and legitimate process executions in the HIS data.

The stakeholders at the hospital whom I interviewed for workaround identification were the Privacy Officer, the CMIO, the CNIO, a system administrator, a physician, and a nurse at an outpatient clinic. This diverse pool of participants gave me multiple perspectives on the processes at the hospital. The privacy officer (who was previously also a nurse) played a central role in the workaround identification. Not only did she come up with several workarounds herself, but she also brought me in contact with the other participants. She was also my main contact at the hospital and the domain expert I could consult for most of my questions.

3.4.2 Event data

The case study hospital's HIS (Chipsoft HiX) is a Process-Aware Information Systems (PAISs) which provides very detailed logs about the activities that have been executed in the system (Mans et al., 2008; van der Aalst, 2016). This event data needed to be manually extracted from the HIS, which was done in accordance with the principle of data minimization so that only data was collected that was necessary for detecting the workarounds. Due to the iterative improvement of the workaround detection, this meant that data extraction needed to be repeated several times. The data was extracted from the HIS's database using SQL queries. The data was stored and analyzed on a dedicated server provided by Intermax for the duration of the case study so that it did not leave their secure network and could easily be accessed remotely. Only pseudonymized data was extracted from the HIS.

I opted for January and February of 2020 as the scope of the data for several reasons. The data is recent, yet it is minimally influenced by the COVID-19 outbreak in the Netherlands. This slice of the data contains thousands of instances for each workaround which should be more than enough to create a proof of concept of the workaround detection. I also believe that it captures all the possible variability in the process executions that is relevant. The main reason for choosing this scope was to reduce the amount of data that needed to be processed. For each working day more than one million events are logged, which can become demanding of the server's resources very quickly.

After extracting the event data, I first pre-processed it by transforming the data into a usable format, filtering out test patients, relating sub-specialisms to main specialisms, and relating patients to their aliases in case of duplicate registrations. Then, I wrote custom scripts using the R programming language for statistical computing (R Core Team, 2013) to assign the events to cases in order to construct event logs. This assignment was done based on characteristics such as the timestamp, patient number, employee number, employee function, and the employee specialism. The conditions for these characteristics varied for each type of event and the situation of the patient at that moment. To analyze the event logs, I used the process mining tool Disco by Fluxicon (<https://fluxicon.com/disco>) as it is lightweight and intuitive to use, yet it contains all the relevant features.

3.4.3 Focus group

At the end of the case study, I organized a focus group in order to answer the fourth sub-question. The stakeholders of the hospital who participated in the focus group were the CISO and the privacy officer. Another attendee was the product owner of EPD Insights at NFIR (a sister company of Intermax). He has a lot of experience in designing IT solutions for healthcare and has gotten involved during the research project. I started by presenting my findings throughout the case study, emphasizing the workarounds that I encountered and the degree to which I was able to detect them using process mining. Then, a discussion followed on what should be included in a snapshot of these workarounds in order to initiate process improvement. The focus group was hosted as a video conference using Microsoft Teams and it was recorded in order to better process the results.

3.5 Assessment of Threats to Validity

An analysis of the threats to validity was carried out using the checklist by Wohlin et al. (2012) based on the four categories of threats to validity originally published by Cook & Campbell (1979). Some of the threats assume a classic experimental design where the effect of a treatment on an independent variable is measured for groups of test subjects drawn from a population in an attempt to discover a causal relation. Such threats do not always apply to a design science research project. This

section will cover the relevant threats by Wohlin et al. (2012) and their response for each validity category.

Conclusion validity is concerned with the reproducibility of the research. A form of the threat *fishing for a specific result* is inherent when answering the second sub-question. Different process mining tools and techniques will be applied in an exploratory manner, in the search for a methodology that is able to detect workarounds well. This threat is accepted as it is fundamental for conducting the research as described. When attempting to successfully detect workarounds, the question arises whether there is enough *statistical power*, as well as whether there is *reliability of measures* in order to draw conclusions from the results. To address these matters, sub-question three was formulated in order to theorize how performance should be measured and how false positives and false negatives should be dealt with.

Internal validity asks whether the measured effect is real or whether it was distorted by any confounding factors. An inevitable threat in this category lied in the *instrumentation*. Due to the COVID-19 outbreak in the Netherlands most of the interviews were held over the phone and the focus group was held in a video conference instead of in person. This might have held the participants back from engaging with the subject matter more.

Construct validity deals with whether the object of analysis reflects the subject of the research question. When designing a methodology to successfully detect workarounds, it will be tested for multiple workarounds to avoid being prone to *mono-operation bias*. Furthermore, when checking for the presence of false positives and false negatives, multiple instances of each workaround will be used to avoid the same bias. During the interviews with stakeholders to elicit workarounds, I will make clear to them at the start that it is not an evaluation of how they perform their work, and that using a workaround is not necessarily a bad thing. With this, I aim to mitigate any possible *evaluation apprehension* so that the interviewees do not feel discouraged to tell me about the workarounds that they practice. During the focus groups, I will try my best to stay neutral as much as possible and avoid putting words into the participants' mouths because I want to avoid causing any bias due to *hypothesis guessing* or *experimenter expectancies*.

External validity addresses whether the findings can be generalized to other contexts. When establishing a suitable methodology for detecting the workarounds discovered at the hospital, there is an *interaction of selection and treatment* as the methodology will be aimed at workarounds that are likely to be detectable. Workarounds that are not reflected in the data will be excluded as they will not contribute to establishing the methodology. The chosen case for the case study has little *interaction of setting and treatment* which means that the case is real and representative. Real data is used in the analysis and the HIS which is studied is the market leader in the Netherlands which makes it a common and representative case (Yin, 2014). However, the hospital's culture might differ from other hospitals which could affect the different qualitative parts of the research.

4 Results

In this chapter, the results of the four sub-questions that followed from the case study will be presented. It first covers the alternative goal workarounds that I discovered in the case study organisation from interviews with stakeholders. Then, I will explain my methodology for developing the detection of these workarounds. After that, I will explain how the performance of the workaround detection is measured and report the performance of the studied workarounds. Finally, I will discuss the findings of the focus group on what information should be captured about the workarounds in order to initiate process improvement.

4.1 Alternative Goal Workarounds

From the interviews with stakeholders at the case study hospital, a total of five alternative goal workarounds were discovered. They can be divided into two preliminary themes that describe the nature of the designed path that is used to reach an alternative goal. Figure 5 depicts the discovered workarounds and their classification.

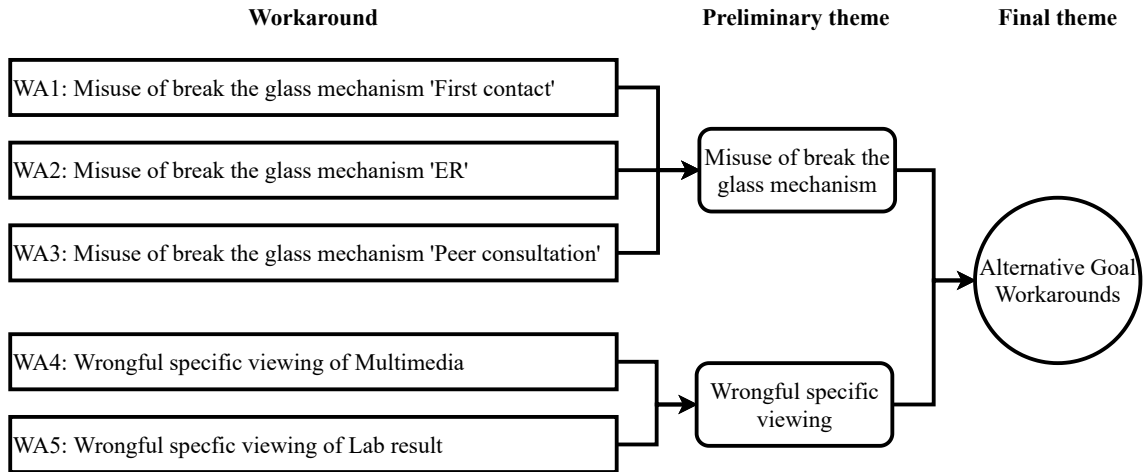


Figure 5: Classification of the discovered workarounds

The first preliminary theme is the misuse of the break the glass mechanism (shown in Figure 6). At the case study hospital, this mechanism is not just used incidentally, but hundreds of times a day. This makes it infeasible to investigate all of the times it was used to verify whether its use was rightful. The current approach for catching misuse is that the privacy officer investigates a random sample of cases on a regular basis. However, because there are far more cases where the mechanism is used rightfully than wrongfully, this approach has been ineffective. If a large portion of cases could automatically be determined to be rightful, the privacy officer would only need to investigate the remainder of the cases and the investigation would

become more effective at catching misuse. This notion sparked the interest in a quantitative workaround detection mechanism at the case study hospital.



Figure 6: 'Break the glass' prompt

When an employee uses the break the glass mechanism, he can enter his reason for doing so either by choosing one of the default options or by entering the reason in the free text field at the bottom. The first three workarounds address the most often used reasons for breaking the glass. By checking whether the rest of the designed path was followed it can be determined whether a case is a workaround or not. For instance, when the reason for breaking the glass is 'ER' (Emergency Room), one of the main signs that the designed path was followed all the way to the intended goal is that the patient was registered at the ER shortly afterwards.

The two remaining workarounds fit the preliminary theme of the wrongful specific viewing. This means that certain modules of the EHR hold information which is sensitive by nature. These modules should be monitored extra carefully as they are at higher risk to be viewed by an unauthorized person. We can narrow down the investigation by finding out when it is necessary for an employee to view the module in order to carry out his job, such as when the employee is treating the patient.

From the interviews I also discovered a number of workarounds that did fit the existing definitions of the term 'workaround'. As there was no conceivable way to systematically detect these workarounds using process mining, the focus of the case study naturally shifted towards the alternative goal workarounds. An overview of these 'classic' workarounds I encountered at the case study hospital is included in Appendix C.

4.2 The PM²4AGW methodology

I have based my methodology for detecting alternative goal workarounds on the Process Mining Project Methodology (PM²) by van Eck et al. (2015). There are several reasons why this methodology in particular forms a good basis for detecting alternative goal workarounds. One of the main goals for process mining projects with PM² is to improve compliance to rules and regulations, which fits well with the notion of making sure that the designed path is only used to reach its intended goal instead of being misused to reach alternative goals. It also emphasizes iterative analysis and the collaboration between process analysts and business experts, both of which have proven to be essential during the case study.

There are however some points at which I deviate from PM². I have formalized this into my adaptation of PM² that I call the Process Mining Project Methodology For Alternative Goal Workarounds, or PM²4AGW for short. The most fundamental deviation is that instead of aiming to answer a number of research questions, PM²4AGW aims to develop and optimize the detection of a workaround. In PM², the iterations serve to refine and add research questions, whereas in PM²4AGW the iterations serve to further optimize the detection based on the evaluation of the previous iteration. This evaluation is based on the findings of the analysis carried out by the business expert, which is a new addition in PM²4AGW. The business expert analyzes a random sample of the cases that were classified as a workaround. This analysis gives an indication of the detection's performance, and insights for further refinement in order to reduce the number of false positives. For clarification on this concept, Figure 7 pictures the confusion matrix of the workaround detection to define the classification dimensions.

		Detected as workaround?	
		<i>Yes</i>	<i>No</i>
Actual workaround?	<i>Yes</i>	True Positive	False Negative
	<i>No</i>	False Positive	True Negative

Figure 7: Confusion matrix for the workaround detection

The PM²4AGW methodology is visualized in Figure 8. Deviations from the original PM² methodology are depicted with a dashed red outline. During the case study, the role of process analyst was fulfilled by me, and the role of business expert by the privacy officer at the hospital.

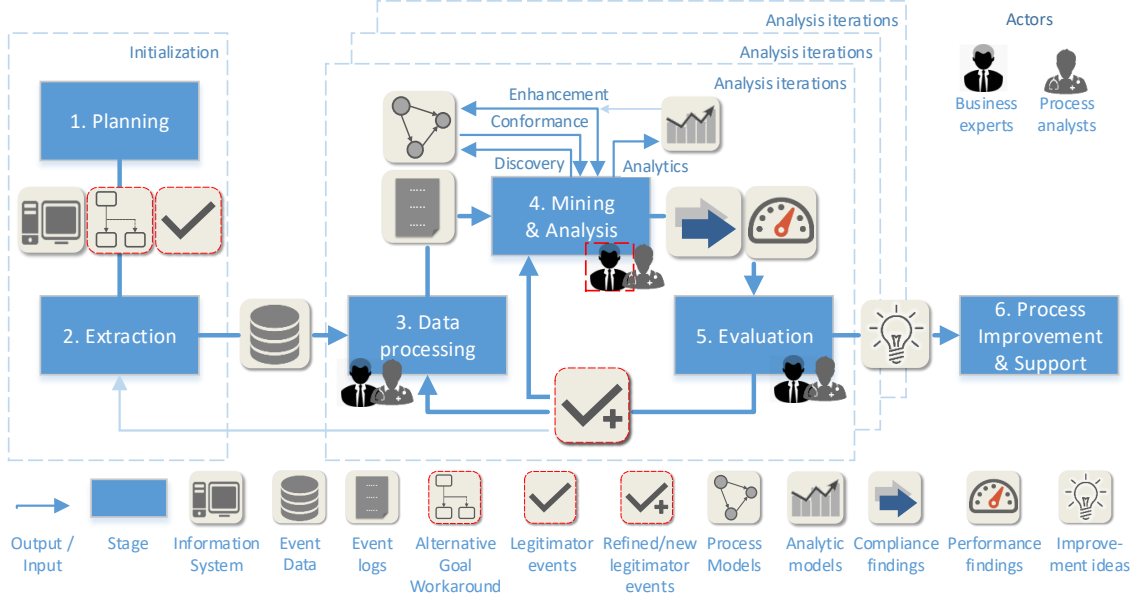


Figure 8: The PM²4AGW methodology (after PM² by van Eck et al., 2015)

1. Planning

The first objective of the planning stage is to define the alternative goal workaround that will be investigated. The workaround is centered around an event class that represents the part of the designed path that is exploited in the workaround. I call this event class the central workaround event, and it must be performed in the execution of every single process instance. When a central workaround event is logged by the system, a case is opened. To determine whether the case is a workaround or a legitimate process instance, legitimitor events are needed. A legitimitor event is an event that occurs when the designed path is followed to the intended goal (instead of the alternative goal). When it occurs in the log of a case, it can serve to legitimize the case. Defining the workaround, the central workaround event, and an initial set of legitimitor events is the starting point for the process mining project.

2. Extraction

In the second stage, data is extracted from the information system for the central workaround events and the legitimitor events that are known thus far. This is one of the main deviations from the PM² methodology, as the event classes are already defined in order to do data extraction in a more targeted fashion. With every iteration new legitimitor events are added, which may require data that has not been extracted yet. Therefore, the extraction stage in PM²4AGW may be repeated more often than is typical in PM².

3. Data processing

During the data processing stage, the extracted data is turned into an event log by constructing cases. The case notion is that every case contains exactly one central workaround event and any number of legitimator events (zero in the case of a workaround). A legitimator event is added to a case if it meets all the conditions specified by the business expert. I implemented the mechanisms for constructing cases by writing custom scripts using the R programming language for statistical computing (R Core Team, 2013).

4. Mining & Analysis




In the mining & analysis stage, the event log is loaded into a process mining tool in order to explore and analyze the process. Conformance checking is carried out a bit differently than usual. Instead of using a reference process model that describes the intended behavior, we simply look at the case variants. When a case contains only the central workaround event (in other words, no legitimator events) that case is classified as a workaround. Additionally, a random sample of ten workarounds cases is drawn which is analyzed by the business expert. By determining which cases are indeed workarounds and which are false positives, the business expert provides input for the evaluation stage.

5. Evaluation

The evaluation stage concludes the analysis iteration and transitions into either the next analysis iteration or the final stage of the methodology. When any workaround cases have been determined to be false positives by the business expert, they will provide input for the next iteration. In collaboration with the business expert, new legitimator events are defined (or existing ones are refined) that will prevent these cases and equivalent ones from being classified as a workaround in the next iteration. This helps to reduce the number of false positives with each iteration in a structured manner.

6. Process Improvement & Support

The final stage of the methodology is reached when the process analyst and the business expert agree that the workaround detection is mature enough. For example, this might be when the last iteration concludes without encountering any false positives in the analysis and thus, there is no input for a new iteration. When fully optimized, the workaround detection can be implemented for continuous operational support to detect workarounds in real-time. An increased awareness of the workaround can also be an incentive for process improvement, as will be addressed in the fourth sub-question.

Case ID	Timestamp	Activity	Class
1	10:04	Break the glass mechanism 'First contact' used	
	10:07	Appointment with patient made	
2	11:42	Break the glass mechanism 'First contact' used	
3	12:35	Referral of patient registered	
	15:49	Break the glass mechanism 'First contact' used	



	Legitimate case
	Workaround case

Figure 9: Simplified example of event log for WA1

To provide a better understanding of how the workaround detection operates in practice, I will provide some concrete examples. Figure 9 visualizes a simplified event log of workaround 1, misuse of break the glass mechanism 'First contact'. The event log consists of three cases that each contain one or multiple activities. The first case shows that shortly after the central workaround event, *Break the glass mechanism 'First contact' used*, a legitimator event is executed, namely *Appointment with patient made*. Because a legitimator event can be related to the central workaround event, the case is legitimate, as shown in the Class column. In case 2, only the central workaround event was executed. With no way to legitimize this event, it is labeled as a workaround case. The last case shows that some legitimator events, in this case *Referral of patient registered*, can also occur before the central workaround event, as long as they fall within the specified time window. Important to note is that in reality there are many more conditions for matching the legitimator events to the central workaround events. These conditions look at various characteristics regarding both the patient and the employee who performs the activity. An example is that some legitimator events only holds if the employee has a certain function at the department where the patient is being treated.

4.3 Measuring Workaround Detection Performance

This section will cover performance measures for the five alternative goal workarounds that I detected during the case study and how they improved with each iteration. When looking at the performance, there are two key figures to consider. The first one is the *percentage of workarounds* that are detected in the data set, which is defined as follows:

$$percentage\ of\ workarounds = \frac{workaround\ cases}{total\ cases} \times 100\%$$

This number alone does not necessarily show whether the detection mechanism has been optimized or not, because the share of true positives is unknown. Some

workarounds might just occur a lot more often in the data set than others, meaning that their percentage will be higher even when fully optimized. This also makes it difficult to compare the workarounds to each other. What does tell us something about the optimization of the workaround detection, is the progression of the percentage of workarounds over time. As the workaround detection matures, the curve levels out, which means that most of the false positives have been eliminated, as far as possible. The percentage of workarounds across the iterations is shown in Figure 10. Note that the y-axis is truncated at 60% to better show the differences between the data points. It would not make sense to compare the workarounds in the same iteration number, as the iterations are not aligned over time. I started with developing the detection of workaround 1 and 2, then workaround 3 followed one iteration later, and workaround 4 and 5 another iteration later. A consequence of this is that some workarounds build forward on knowledge from earlier iterations of other workarounds. For example, the first iteration of workarounds 3, 4, and 5 use a mechanism that was first discovered during the second iteration of workaround 1. Because of this time discrepancy, workarounds that started at a later point in time may have date points that are lower than they otherwise would be.

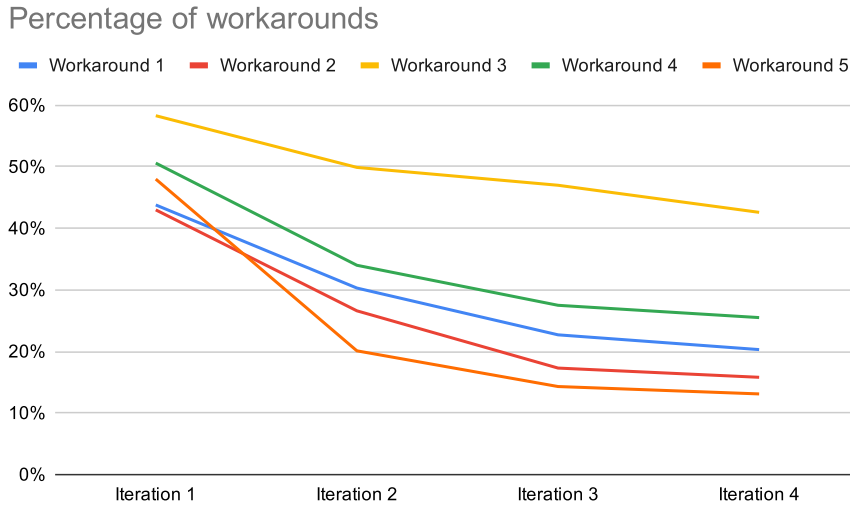


Figure 10: The percentage of workarounds across the iterations

The second key figure is the *precision score* of the workaround detection which is estimated from the small random sample of cases analyzed by the business expert. It is calculated as follows:

$$\text{precision score} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

The precision score tells us more about the maturity of the workaround detection as the business expert investigated a number of cases in order to determine whether

they are actually workarounds or not. As the curve approaches the maximum precision score, there are very little false positives left which also means little input for improvement. This is the biggest indicator that the workaround detection has nearly been optimized. The precision score of the workarounds across the iterations is shown in Figure 11. I used a sample size of ten for estimating the precision score during the case study. If the resources allow for it, a larger sample size would create a better reflection of the workaround cases. On the other hand, a smaller sample size would not be advisable. As a sample size of ten is relatively small, the scores are expected to be impacted by sampling error quite a bit. For example, it looks as though workaround 3 did not improve during the second iteration, as the precision score remains the same. However, this may be skewed by the sample of the first iteration having an over-representation of true positives. The first ground for this suspicion is that 0.7 is an unusually high precision score for the first iteration, when comparing it to the other workarounds. Secondly, in Figure 10 it can be observed that the percentage of workarounds decreased by nearly 10%. In other words, a lot of false positives were eliminated so there was in fact a relatively large improvement in the second iteration.

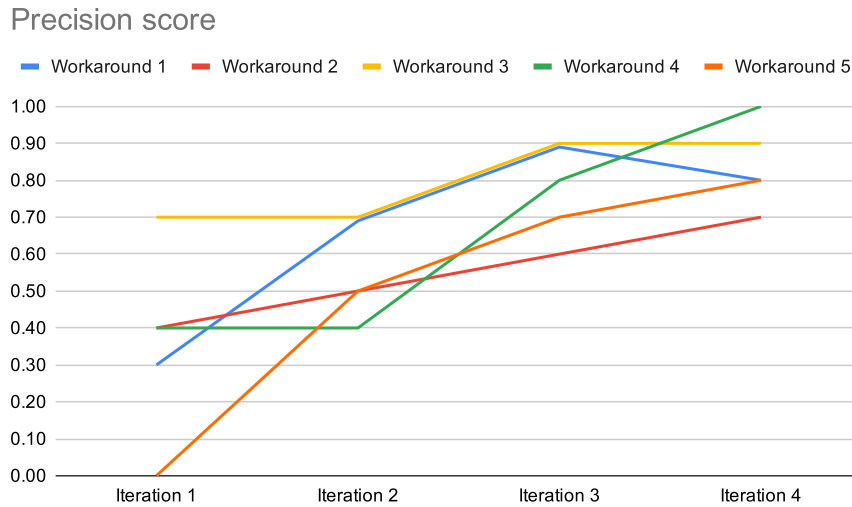


Figure 11: The precision score of the workarounds across the iterations

There is a third figure which may be considered of interest, which is the *number of legitimator events* used in the workaround detection mechanism. It does not tell us anything directly about the performance but rather about the maturity and complexity of the workaround detection. The number of legitimator events is shown for each workaround across the iterations in Figure 12. Important to note is that there are large differences in the effectiveness of legitimator events in reducing false positives. Some legitimator events are very general and some are only applicable

to a specific department of the hospital. Another thing that this number does not reflect is when legitimator events are refined. Even though new false positives are eliminated which does cause an increase in performance, no new legitimator event are added to the count. Because of these shortcomings, it is difficult to draw comparisons between the workarounds and between iterations. In general, the figure shows how the development of new legitimator events slows down over time as less and less input is provided for new iterations.

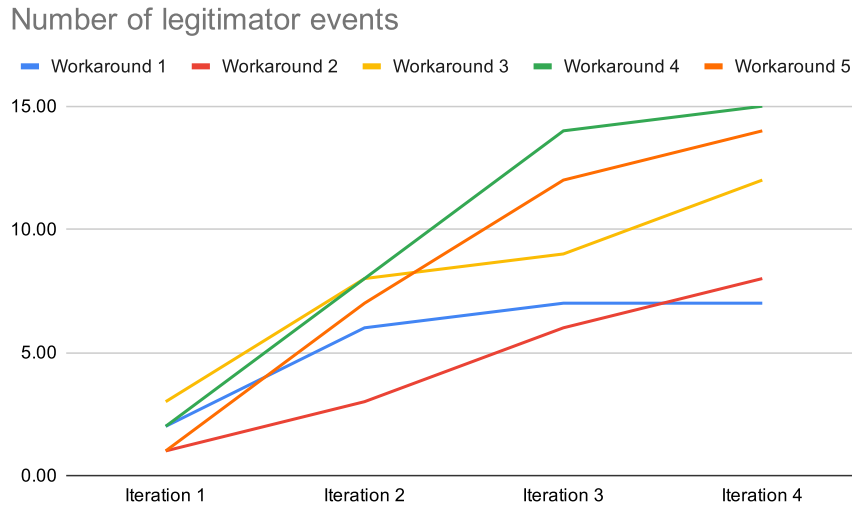


Figure 12: The number of legitimator events for the workarounds across the iterations

So far, I have only covered the reduction of false positives and not false negatives. The main reason why the PM²4AGW methodology puts the emphasis on false positives is that they are simply easier to detect than false negatives. By analyzing sample cases that have been classified as a workaround, the business expert can easily determine whether the case is truly a workaround, or a false positive and why. This directly leads to new or refined legitimator events that reduce the number of false positives. False negatives are harder to identify, as they remain hidden in the larger portion of the cases that has already been systematically determined to be legitimate due to one or multiple legitimator events. When analyzing a sample of workaround cases, you are almost certain to encounter false positives, especially during the earlier iterations. Whereas when analyzing a sample of non-workaround cases, it is less likely that you will encounter false negatives. The reason for this is that there are far more true negatives than true positives. In other words, the data is skewed. Another reason lies in the origin of false negatives. When the process analyst and the business expert work together meticulously, very little false negatives should ensue. A false negative will either be caused by a legitimator event that has

been specified too broadly or due to an implementation error. To counteract the first risk, it is recommended that the legitimator event specifications are critically peer reviewed by a second business expert. The second possible origin of false negatives is implementation errors that may cause workarounds to fly under the radar, which is more challenging to localize.

By analyzing a random sample of non-workaround cases, false negatives can be detected in a way similar to how false positives are detected at the end of each analysis iteration. However, as mentioned before, this method will be less effective at detecting false negatives than false positives. In reality, the business expert will encounter a lot of known case variants that have already been determined to be legitimate based on the previously specified legitimator events. For an organization that applies the PM²4AGW methodology to develop workaround detection for operational support, the decision of focusing on reducing false positives or false negatives is also a matter of resource allocation. As detecting false positives is relatively easier it is also more impactful in improving the workaround detection. Therefore, it is likely to receive more resources, especially in the early stages of development when there is still a large number of false positives and little to no false negatives.

4.4 Snapshot for AGWs

Workaround snapshots have previously been used by Beerepoot & van de Weerd (2018) to capture knowledge about workarounds in order to understand the users' needs and how the HIS does not fulfill them, leading to the emergence of the workarounds. For alternative goal workarounds we want to achieve a similar goal, to capture knowledge about the workarounds in order to initiate process improvement. However, some of the existing components of the workaround snapshot in Beerepoot & van de Weerd (2018) are not applicable to AGWs in their current form.

For example, it may be difficult to name the roles of the *workers* who practice the workarounds. The alternative goal workarounds discovered in the case study hospital are quite general and could be carried out by almost every role. It is also challenging to provide a complete *description* of the workaround as the alternative goal that the workaround tries to achieve can vary between instances. What could be useful is to describe the central workaround event that is used to reach the alternative goal in the workaround. It would not make a lot of sense to create a *process model* of an alternative goal workaround. The workaround case variant consists only of the central workaround event. On the other hand, the large amount of known legitimate case variants may consist of any combination of the legitimator events. There is no relevant logic captured about the (co-)occurrence of the different legitimator events. What would be more useful is to create a list of all the legitimator events of the workaround. Both the *impact* and the *motivation* of the workaround are unknown and will differ across different alternative goals that are reached with the workaround.

Because the workaround snapshot by Beerepoot & van de Weerd (2018) is not directly applicable to alternative goal workarounds, a focus group was organized with a number of stakeholders to discuss how the workaround snapshot should be given shape for AGWs. The attendees of the focus group were the hospital’s CISO and privacy officer, and the product owner of EPD Insights at NFIR. First, the focus group came to a consensus that next to using the workaround detection to support the privacy officer’s job by alerting her of cases that have a high probability of being workarounds, knowledge about the workarounds can also be used to initiate process improvement. Whereas detection is used to catch cases where workarounds are used to reach malicious alternative goals, process improvement aims to address cases where they are used to reach non-malicious alternative goals. To enable process improvement, the snapshot should contain the components shown in Table 1.

Snapshot component	Content
Date of snapshot	Date the snapshot was created
Central workaround event	Definition of the central workaround event
Legitiminator events	List of defined legitimator events
Degree of occurrence	How often the central workaround event occurs
Percentage of workarounds	The share of cases that are workarounds
Precision score	Estimate of the precision score in workaround cases
Workers in workaround cases	Overview of roles involved in workaround cases
Workers in legitimized cases	Overview of roles involved in legitimized cases

Table 1: Snapshot components for alternative goal workarounds

As mentioned before, the description and process model are replaced by the *central workaround event* and the *legitiminator events*. These describe what part of the designed path is misused and what characteristics have been considered to distinguish process instances that reach the intended goal from those that reach an alternative goal. The *degree of occurrence*, *percentage of workarounds*, and the *precision score* help with estimating the impact of the workaround, as well as the maturity of the detection. By creating an overview of both the *workers in workaround cases* and the *workers in legitimized cases*, it can be observed whether these are the same groups of people or not.

In order to initiate process improvement with the snapshot, we first need to start a dialogue. Some workers of both workaround cases and legitimized cases are invited to discuss how they perform their work and to find out what makes their cases to be considered a workaround or legitimate. Then the organization can further discuss whether there needs to be a change in the system or in the way of working to address the non-malicious workaround cases. This enables the organization to not only improve the process, but also to distinguish between malicious and non-malicious goals, further empowering the privacy officer to catch malicious behavior.

5 Discussion

The research presented in this thesis can be positioned in a number of ways. This study broadens the scope of research on workarounds by defining and exploring a new type of workaround, the alternative goal workaround (AGW). I also contribute to filling a research gap as this is only the second study that attempts to detect workarounds quantitatively (by using process mining), joining the study by Outmazgin & Soffer (2013). Another way that this study can be positioned is as part of the class of research that applies process mining in the healthcare domain. Finally, it also presents a new application of the workaround snapshot by Beerepoot & van de Weerd (2018) adapted for AGWs in order to initiate process improvement.

This study differentiates itself from the study by Outmazgin & Soffer (2013) in four main aspects. First of all, Outmazgin & Soffer used a number of 'classic' workarounds, whereas I studied alternative goal workarounds which are defined differently. Secondly, Outmazgin & Soffer studied purchasing processes, whereas I studied healthcare processes which have been characterized as highly complex (Mans et al., 2008). Third, Outmazgin & Soffer focused on detecting generic workarounds without requiring additional domain knowledge. This is vastly different from the workarounds that I detected, as they did require gathering additional domain knowledge about the processes with each iterative cycle. Finally, the way in which process mining was applied to detect the workarounds was fundamentally different. Outmazgin & Soffer used filters that looked for pre-defined log patterns in their event logs to detect their workarounds. On the other hand, the PM²4AGW methodology was designed to fit the detection mechanism to the specific context of a workaround and its main value lays in constructing the event log itself. It starts out with an initial set of legitimator events which results in a high percentage of supposed workarounds in the data set. By analyzing a small sample of these supposed workarounds in collaboration with a business expert, new legitimator events can be defined to prevent the false positives in the sample from re-appearing in the next iterative cycle. As these cycles continue, more and more false positives are filtered from the data set. Meanwhile, the true positives stay behind, increasing the precision score of the workaround detection. This methodology can be explained as the inverse of the approach by Outmazgin & Soffer. Instead of using the pattern of the workaround itself, the patterns of legitimate process instances are used to expose the workarounds. As the studied alternative goal workarounds are too diverse and complex to be captured in a single pattern, the PM²4AGW methodology proved to be an effective solution.

To expand on the discussion in section 4.3 regarding the detection of false positives and false negatives, I suggest two more methods for detecting these errors besides analyzing a random sample of cases. By doing *manual testing*, the business expert can perform behavior in the HIS that should be detected as a workaround to make sure that there are no false negatives. Or in the opposite case, the business expert can perform legitimate behavior to check whether the detection does not

produce any false positives. This testing method is more targeted than analyzing a random sample of cases, but it is also dependent on human ability. It relies on the human to be consistent and complete in the testing. There may be variants of behavior that are unknown to the business expert and its potential false positives or false negatives will remain undetected by this method. To be less reliant on humans there is also the possibility of doing *automated testing*. By simulating the behavior in the HIS using software, the testing can be performed without any human interference. There will be an initial implementation cost, but it will ultimately be more efficient and less prone to human error than manual testing. This could for instance be accomplished by using Robotic Process Automation (van der Aalst et al., 2018). Whereas the detection of false positives is already effectively addressed in PM²4AGW through analyzing random samples, it does not yet include a method to detect false negatives. If the organization also wants to spend resources on detecting false negatives, I recommend manual testing, as it will be more effective than analyzing a random sample of cases but not as costly as automated testing. Automated testing may be considered in a later stage for detecting both false positives and false negatives when the workaround detection has matured and has been fully implemented for operational support. It will help to ensure continuity of the detection mechanism when changes are made to the HIS. These methods for testing quantitative workaround detection could be a starting point for future research.

One of the main limitations of the PM²4AGW methodology is that it was specifically designed for alternative goal workarounds. However, it might inspire future researchers to consider using PM², PM²4AGW, or an adaptation of their own for detecting workarounds from event data with process mining. Another limitation is that the methodology was only demonstrated in a single case study on a total of five AGWs. While it can be argued that the case study hospital is representative, applying the methodology in more case studies and to more AGWs would solidify the findings. Follow-up studies could also look outside the healthcare sector to find out whether AGWs can be discovered and detected in a broader range of information systems and processes.

As was previously pointed out by Koppel et al. (2015), qualitative investigation of the processes is a prerequisite for understanding workarounds in the complex problem domain that is healthcare. On a similar note, van Eck et al. (2015) concluded that process mining is most effective when process analysts work closely together with business experts in a highly iterative and interactive manner. I would also like to endorse this as a recommendation for future researchers in this research area. Without the close collaboration with domain experts it would have been impossible for me to develop workaround detection in these processes, let alone to fully understand them.

6 Conclusion

In this thesis, I explored a new type of workaround where instead of circumventing a designed path to reach the intended goal, the user exploits (a part of) the designed path to reach an alternative goal (either malicious or non-malicious). I narrowed the research gap in quantitative workaround detection by answering the following research question: *"How can process mining be used to detect workarounds that are used to reach an alternative goal?"*.

I discovered five alternative goal workarounds (AGWs) that are practiced at a Dutch hospital, which can be subdivided into two themes. By designing the PM²4AGW methodology, I have found a way to distinguish between workarounds and legitimate process instances using process mining. The successful detection of the workarounds can be attributed to my approach that is based on the principle of exclusion. Instead of looking for needles in a haystack, we develop an efficient way to remove hay from the stack until only the needles remain. Two other critical characteristics of the methodology are the collaboration between process analysts and business experts, and working in iterative cycles.

The workaround detection mechanism that I built can automatically determine for the lion's share of the cases that they are legitimate based on the occurrence of certain events in the HIS. The privacy officer of the hospital now only needs to investigate a small fraction of the cases that remains suspicious, whereas previously she would investigate a sample out of all the cases. Thus, the workaround detection empowers her to more efficiently catch and sanction unlawful use of EHRs, which is a big step forward in assuring the privacy of patients.

Acknowledgements

First of all, I want to express my gratitude to the hospital where I performed the case study. My special thanks go to the privacy officer who helped me tremendously in understanding the processes and with carrying out the many analysis cycles to optimize the workaround detection. I would also like to thank the other stakeholders who took the time to participate in my study.

Secondly, I highly appreciate the help of my three supervisors at Utrecht University. As my main supervisor, Iris contributed greatly in shaping my study the way it turned out. I would also like to thank Inge and Xixi for their enthusiasm towards my research and their helpful suggestions.

Finally, I am grateful for the people at Intermax and NFIR for enabling me to carry out this project under their supervision, in particular René and Sergej.

References

- van der Aalst, W. M. P. (2016). *Process mining: Data science in action*. Springer. DOI: 10.1007/978-3-662-49851-4.
- van der Aalst, W. M. P., Adriansyah, A., Medeiros, A. K. A. de, Arcieri, F., Baier, T., Blickle, T., et al. (2011). Process Mining Manifesto. In: *International Conference on Business Process Management*. Vol. 99, pp. 169–194. DOI: 10.1007/978-3-642-28108-2_19.
- van der Aalst, W. M. P., Bichler, M., & Heinzl, A. (2018). Robotic Process Automation. In: *Business and Information Systems Engineering* 60 (4), pp. 269–272. DOI: 10.1007/s12599-018-0542-4.
- Alter, S. (2014). Theory of workarounds. In: *Communications of the Association for Information Systems* 34 (55), pp. 1041–1066. DOI: 10.17705/1cais.03455.
- Andrade, E., Leopold, H., van der Aa, H., Alter, S., & Reijers, H. A. (2016). Factors leading to business process noncompliance and its positive and negative effects: Empirical insights from a case study. In: *AMCIS 2016: Surfing the IT Innovation Wave - 22nd Americas Conference on Information Systems*.
- Arduin, P. E., & Vieru, D. (2017). Workarounds as means to identify insider threats to information systems security. In: *AMCIS 2017 - America's Conference on Information Systems: A Tradition of Innovation 2017*, pp. 1–5.
- Beerepoot, I., Koorn, J., Weerd, I. van de, Hooff, B. van den, Leopold, H., & Reijers, H. A. (2019). Working Around Health Information Systems: The Role of Power. In: *Proceedings of the Fortieth International Conference on Information Systems, Munich 2019*. Association for Information Systems.
- Beerepoot, I., Ouali, A., Weerd, I. van de, & Reijers, H. A. (2019). Working around health information systems: to accept or not to accept? In: *Twenty-Seventh European Conference on Information Systems*.
- Beerepoot, I., & van de Weerd, I. (2018). Prevent, redesign, adopt or ignore: Improving healthcare using knowledge of workarounds. In: *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*.
- Beerepoot, I., Weerd, I. van de, & Reijers, H. A. (2018). Detecting Workarounds Using Domain Knowledge-driven Process Mining. In: *Thirty ninth International Conference on Information Systems, San Francisco 2018*.
- Borycki, E., Joe, R. S., Armstrong, B., Bellwood, P., & Campbell, R. (2011). Educating health professionals about the electronic health record (EHR): Removing the barriers to adoption. In: *Knowledge Management and E-Learning* 3 (1), pp. 51–62. DOI: 10.34105/j.kmel.2011.03.006.
- Bozkaya, M., Gabriels, J., & van der Werf, J. M. (2009). Process diagnostics: A method based on process mining. In: *Proceedings - International Conference on Information, Process, and Knowledge Management, eKNOW 2009*, pp. 22–27. DOI: 10.1109/eKNOW.2009.29.

- Burgoon, J. K. (1982). Privacy and Communication. In: *Annals of the International Communication Association* 6 (1), pp. 206–249. DOI: 10.1080/23808985.1982.11678499.
- Burns, A. J., Young, J., Courtney, J. F., Roberts, T. L., & Selwyn Ellis, T. (2015). Exploring the Role of Contextual Integrity in Electronic Medical Record (EMR) System Workaround Decisions: An Information Security and Privacy Perspective. In: *AIS Transactions on Human-Computer Interaction* 7 (3), pp. 142–165.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation - Design and Analysis Issues for Field Settings*. Vol. 3. Houghton Mifflin Company.
- Cresswell, K. M., Mozaffar, H., Lee, L., Williams, R., & Sheikh, A. (2017). Workarounds to hospital electronic prescribing systems: A qualitative study in English hospitals. In: *BMJ Quality and Safety* 26 (7), pp. 542–551. DOI: 10.1136/bmjqs-2015-005149.
- van Eck, M. L., Lu, X., Leemans, S. J. J., & van der Aalst, W. M. P. (2015). PM²: A Process Mining Project Methodology. In: *International Conference on Advanced Information Systems Engineering*. Vol. 9097. Springer, pp. 297–313. DOI: 10.1007/978-3-319-19069-3_19.
- Ejnefjäll, T., & Ågerfalk, P. J. (2019). Conceptualizing Workarounds: Meanings and Manifestations in Information Systems Research. In: *Communications of the Association for Information Systems* 45, pp. 340–363. DOI: 10.17705/1cais.04520.
- International Organization for Standardization (1998). *Information Technology - Vocabulary*. Standard ISO/IEC 2382:1998.
- International Organization for Standardization (2011). *Health Informatics - Classification of purposes for processing personal health information*. Standard ISO/TS 14265:2011.
- Kirlappos, I., Parkin, S., & Sasse, M. A. (2014). Learning from “Shadow Security”: Why Understanding Non-Compliant Behaviors Provides the Basis for Effective Security. In: *USEC’14 Workshop on Usable Security*. DOI: 10.14722/usec.2014.23007.
- Kirlappos, I., Parkin, S., & Sasse, M. A. (2015). “Shadow security” as a tool for the learning organization. In: *ACM SIGCAS Computers and Society* 45 (1), pp. 29–37. DOI: 10.1145/2738210.2738216.
- Kobayashi, M., Fussell, S. R., Xiao, Y., & Seagull, F. J. (2005). Work coordination, workflow, and workarounds in a medical context. In: *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1561–1564. DOI: 10.1145/1056808.1056966.
- Koppel, R., Smith, S. W., Blythe, J., & Kothari, V. (2015). Workarounds to Computer Access in Healthcare Organizations: You Want My Password or a Dead Patient? In: *Studies in Health Technology and Informatics* 208, pp. 215–220. DOI: 10.3233/978-1-61499-488-6-215.

- Koppel, R., Wetterneck, T., Telles, J. L., & Karsh, B.-T. (2008). Workarounds to Barcode Medication Administration Systems: Their Occurrences, Causes, and Threats to Patient Safety. In: *Journal of the American Informatic Association* 15, pp. 408–423. DOI: 10.1197/jamia.M2616.
- Leino-Kilpi, H., Välimäki, M., Dassen, T., Gasull, M., Lemonidou, C., Scott, A., & Arndt, M. (2001). Privacy: A review of the literature. In: *International Journal of Nursing Studies* 38 (6), pp. 663–671. DOI: 10.1016/S0020-7489(00)00111-5.
- Lovis, C., Spahni, S., Cassoni, N., & Geissbuhler, A. (2007). Comprehensive management of the access to the electronic patient record: Towards trans-institutional networks. In: *International Journal of Medical Informatics* 76 (5-6), pp. 466–470. DOI: 10.1016/j.ijmedinf.2006.09.014.
- Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M., & Bakker, P. J. (2008). Application of process mining in healthcare - A case study in a Dutch Hospital. In: *Communications in Computer and Information Science* 25, pp. 425–438. DOI: 10.1007/978-3-540-92219-3_32.
- Mans, R. S., van der Aalst, W. M. P., & Vanwersch, R. J. B. (2015). *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Springer. DOI: 10.1007/978-3-319-16071-9.
- Murphy, A. R., Reddy, M. C., & Xu, H. (2014). Privacy practices in collaborative environments: A study of emergency department staff. In: *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, pp. 269–282. DOI: 10.1145/2531602.2531643.
- Outmazgin, N. (2013). Exploring Workaround Situations in Business Processes. In: *International Conference on Business Process Management*. Springer, pp. 426–437. DOI: 10.1007/978-3-642-36285-9_45.
- Outmazgin, N., & Soffer, P. (2013). Business process workarounds: What can and cannot be detected by process mining. In: *Lecture Notes in Business Information Processing*, pp. 48–62. DOI: 10.1007/978-3-642-38484-4_5.
- Peffer, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. In: *Journal of Management Information Systems* 24 (3), pp. 45–77. DOI: 10.2753/MIS0742-1222240302.
- Ramezani, E., Fahland, D., & van der Aalst, W. M. (2012). Where did I misbehave? Diagnostic information in compliance checking. In: *International conference on business process management*. Springer, pp. 262–278.
- R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- Rebege, Á., & Ferreira, D. R. (2012). Business process analysis in healthcare environments: A methodology based on process mining. In: *Information Systems* 37 (2), pp. 99–116. DOI: 10.1016/j.is.2011.01.003.

- Röder, N., Wiesche, M., Schermann, M., & Krcmar, H. (2014). Why managers tolerate workarounds - The role of information systems. In: *20th Americas Conference on Information Systems, AMCIS 2014*.
- Röder, N., Wiesche, M., Schermann, M., & Krcmar, H. (2016). Toward an ontology of workarounds: A literature review on existing concepts. In: *49th Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 5177–5186. DOI: 10.1109/HICSS.2016.640.
- Rodrigues, J. J., De La Torre, I., Fernández, G., & López-Coronado, M. (2013). Analysis of the security and privacy requirements of cloud-based electronic health records systems. In: *Journal of Medical Internet Research* 15 (8), pp. 1–10. DOI: 10.2196/jmir.2494.
- Rojas, E., Munoz-Gama, J., Sepúlveda, M., & Capurro, D. (2016). Process mining in healthcare: A literature review. In: *Journal of Biomedical Informatics* 61, pp. 224–236. DOI: 10.1016/j.jbi.2016.04.007.
- Scholz, R. W., & Tietje, O. (2002). *Embedded Case Study Methods: Integrating quantitative and qualitative knowledge*. Sage Publications.
- Smith, S. W., & Koppel, R. (2014). Healthcare information technology’s relativity problems: A typology of how patients’ physical reality, clinicians’ mental models, and healthcare information technology differ. In: *Journal of the American Medical Informatics Association* 21 (1), pp. 117–131. DOI: 10.1136/amiajnl-2012-001419.
- Vogelsmeier, A. A., Halbesleben, J. R. B., & Scott-Cawiezell, J. R. (2008). Technology Implementation and Workarounds in the Nursing Home. In: *Journal of the American Medical Informatics Association* 15 (1), pp. 114–119. DOI: 10.1197/jamia.M2378.
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in software engineering*. Springer Science & Business Media. DOI: 10.1007/978-3-642-29044-2.
- Yin, R. K. (2014). *Case Study Research: Designs and Methods*. 5th ed. Sage Publications.

A Information security practices

Like all hospitals in the Netherlands, the case study hospital is required to comply with the various laws and regulations that apply to providing medical care. These include but are not limited to the ones already mentioned in the literature section on patient privacy, namely the GDPR and the NEN 7510 with extensions. Additionally, they mentioned needing to comply with the WGBO, which is the Dutch Medical Treatment Contracts Act and guidelines from the KNMG, the Royal Dutch Medical Association.

At the case study hospital, they define privacy as *"the right to be left alone"*. What is and is not sufficient to guarantee privacy is determined by the aforementioned laws and regulations. What differentiates Dutch hospitals from each other, is the package of technical and organisational measures for information security and privacy that they develop in order to comply with the laws and regulations. These measures are captured in a bundle of documented policies that are carried out in the organisation from the top down. Now an overview will follow of the policies at the case study hospital.

First of all, employees are bound to a code of conduct, which contains instructions on how to handle sensitive information such as what may or may not be shared through certain channels. It also includes a passage on the need for de-authentication, i.e., lock your workstation when you leave it unattended. Employees are encouraged to address colleagues who do not adhere to this policy. Computers also lock automatically after 10 minutes of inactivity. The hospital uses two-factor authentication by combining a physical pass and a login code.

This brings us to the authorization policy, which assumes proper authentication, i.e., that users only perform work using their own identity. Authorization for accessing a patient's record should be limited to employees who have a treatment relationship with the patient. Additionally, the employees should only have access to the parts of the record that are required for them to carry out their job. It has been found to be infeasible to orchestrate this on an individual level, as there are hundreds of outpatient clinic visits each day. Instead, the hospital uses authorization profiles which are based on a user's role and medical specialism. The hospital employs an authorization expert whose task is to maintain and assign the authorization profiles.

A limitation on the authorization of an employee should never be the cause of failing to deliver the necessary care to a patient. Therefore, the hospital employs the 'break the glass'-mechanism as covered in section 2.5. Employees can 'break the glass' to gain access to most of the patient records that normally would be prohibited for them. When they attempt to access such a record, a message pops up that informs them that all their actions are logged and unlawful use will be punished. They can enter their reason for breaking the glass and continue to the record. There are some exceptions, such as some particularly sensitive records (e.g. psychiatry) which are blocked for employees of different medical specialisms.

Furthermore, there is a privacy declaration for patients, and one for employees, which explain in detail what data is captured about those people and for what purpose it is used.

To inform new employees of the policies, their introduction meeting includes a presentation on information security and privacy. The hospital has also made it mandatory for all employees to complete an e-learning module on information security and privacy. This module teaches them how to handle different situations that will be relevant to them, based on their role at the hospital. New employees must complete the module within three months of entering employment. There is also a frequently asked questions page to help employees as it can sometimes become confusing what they are and are not allowed to do, especially for new employees.

There are documented procedures for doing regular checks to ensure compliance with the policies. One check is concerned with assessing the level of information security maturity on the work floor. Another check selects a number of random patient records to thoroughly investigate whether there has been any unlawful use. The 'break the glass' mechanism is also checked in this manner, by regularly investigating a number of random cases and checking whether the reason for its use was valid. There is also an external audit that is carried out by an accountancy firm once a year. This is also required for obtaining certification of the NEN 7510 and the ISO 27001 standard for information security management.

When a data leak does occur in the Netherlands, an organisation needs to report it to the AP, the Dutch Data Protection Authority, in order to comply with the WMD, which is the Dutch Data Leak Reporting Obligation Act. The case study hospital has created an internal procedure for what to do when a potential data leak is encountered. Employees must report the leak to the Data Protection Officer right away after which a commission will start investigating the data leak. If the leak is real, the Data Protection Officer will report it to the AP. The hospital also has a policy for how to sanction carelessness and misconduct by employees.

B Contact moments with stakeholders

Date	Stakeholder(s)	Activity	Duration
05-03-2020	CISO & Privacy Officer	Interview on information security practices	1 hr
05-03-2020	Privacy Officer	Introduction of HIS and work observation	4 hrs
16-03-2020	Privacy Officer	Discussing data sources and potential workarounds	1 hr
27-03-2020	Privacy Officer	Discussing data sources and potential workarounds	30 mins
03-04-2020	Privacy Officer	Iterative analysis cycles	1 hr 45 mins
09-04-2020	System Administrator	Interview on potential workarounds	30 mins
24-04-2020	Privacy Officer	Iterative analysis cycles	1 hr
07-05-2020	Privacy Officer	Iterative analysis cycles	1 hr 30 mins
13-05-2020	Nurse	Interview on potential workarounds	30 mins
14-05-2020	Privacy Officer	Iterative analysis cycles	1 hr
19-05-2020	CMIO	Interview on potential workarounds	30 mins
22-05-2020	CNIO & Nurse	Interview on potential workarounds	1 hr
27-05-2020	Privacy Officer	Iterative analysis cycles	1 hr
05-06-2020	Privacy Officer	Iterative analysis cycles	1 hr
17-06-2020	Privacy Officer	Iterative analysis cycles	1 hr
26-06-2020	Privacy Officer	Iterative analysis cycles	1 hr 30 mins
07-07-2020	CISO & Privacy Officer*	Focus group 'workarounds for process improvement'	1 hr
10-07-2020	Privacy Officer	Iterative analysis cycles	1 hr
17-07-2020	Privacy Officer	Iterative analysis cycles	45 mins
27-07-2020	Privacy Officer	Iterative analysis cycles	1 hr 15 mins
05-08-2020	Privacy Officer	Iterative analysis cycles	30 mins

Table 2: Contact moments with stakeholders at the case study hospital

** This activity also included the Product Owner of EPD Insights, who does not work at the case study hospital*

C Classic workarounds

This appendix reports the miscellaneous 'classic' workarounds that I encountered at the case study hospital. They are classic in the sense that they fit existing definitions of the term 'workaround', as covered in section 2.1. The discovered workarounds are summed up below, preceded by their classification(s) following the overview by Röder et al. (2016):

- **Shadow System:** Certain departments use Excel sheets for planning their work instead of the planning module in the HIS, as this module does not meet their preferences.
- **System Misuse & Non conformity:** Some physicians write the logistical planning regarding a patient in a free text field instead of using the order module in the HIS. Secretaries then need to read through this text field to figure out what needs to happen with a patient.
- **Tweaking & System Misuse:** Incoming patients at the ER get a temporary registration in order to monitor the expected workload. This temporary registration is then deleted instead of being linked to the complete registration.
- **Reinvention:** There is a 'reason for discharge' that employees can select that has been set up in a way to circumvent certain restrictions of the HIS.