UHASSELT

KNOWLEDGE IN ACTION

Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Establishing requirements for the application of process discovery algorithms in audit practice

Elseline Senave

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

Prof. dr. Mieke JANS

BEGELEIDER :

Mevrouw Manal LAGHMOUCH

UHASSELT

KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

2020
2021

Faculteit Bedrijfseconomische Wetenschappen

master handelsingenieur in de beleidsinformatica

Masterthesis

Establishing requirements for the application of process discovery algorithms in audit practice

Elseline Senave

Scriptie ingediend tot het behalen van de graad van master handelsingenieur in de beleidsinformatica

PROMOTOR :

Prof. dr. Mieke JANS

BEGELEIDER :

Mevrouw Manal LAGHMOUCH

This master thesis was written during the COVID-19 crisis in 2020-2021. This global health crisis might have had an impact on the (writing) process, the research activities and the research results that are at the basis of this thesis.

Preface

"Feel the wind start to change, take me through the night. Now I'm not far away from the brightest side." (Katelyn Tarver, 'Somebody Else')

After a year that seemed to last forever and insisted on leaving me guessing, I am proud to finally send the final version of my master thesis out into the world. Before you lies 'Establishing requirements for the application of process discovery algorithms in audit practice', a paper through which I endeavoured to provide the means for assessing process discovery algorithms on their usefulness in the next chapter of audit practice, Audit 2.0. This study could have never been realised without the help and support of the following people, to whom I would love to express my gratitude in this preface.

First and foremost, thank you Manal Laghmouch. Over the last year, you went above and beyond to help me reach my full potential. Thank you for your valuable advice, for the many read-throughs and feedback sessions, and for always being available and approachable for questions and assistance. Your contributions lifted this work up to heights I could have never reached on my own.

Thank you Professor Jans, for sharing your knowledge and experience. Your considerate, reassuring and encouraging approach allowed the creation of this thesis to be not only an instructive, but also an enjoyable experience. Thank you for leading the way.

Thirdly, Anjulie Senave. My sister, my best friend, my all-round favourite person. Thank you for the helpful and uplifting phone calls, for taking on the auditor perspective, for aiding me with the final spellcheck, and for always being on my team, even when I sometimes make it difficult to be.

Karel-Jan Senave, thank you for driving all the way down from Antwerp to lend me your laptop when mine refused to do the analyses required for the demonstration. Thank you for being the coolest and most supportive big brother any girl could have, and for sending me cute animal videos whenever I needed a little distraction.

And last but not least, I am indebted to you, Skye, the sun, moon, and stars of my life. Thank you for brightening up my dullest and darkest of days. I could not have done this without you.

From my hands to yours,

Elseline Senave
(June 2021)

Establishing requirements for the application of process discovery algorithms in audit practice

Elseline Senave, Manal Laghmouch and Mieke Jans

Hasselt University, 3590 Diepenbeek, Belgium

Abstract

Whispers of Audit 2.0, a profound reform of audit practice led by the adoption of process mining techniques and tools, carry with them the promise of audit procedures of increased quality, rigidity and efficiency. Namely process mining's flagship approach, process discovery, powers the reconstruction of a company's processes as they transpired in practice, instead of reflecting anyone's expectations of their execution. In response to some of the limitations of traditional audit procedures, the reconstruction would be based on the company's full collection of transactions, as recorded (partly) independently of the auditee in his information systems. However, for audit practice to fully benefit from process discovery's potential, it should be ensured that discovery algorithms play into audit concerns. Hence, this study endeavours to obtain insight into the current perspective on requirements for successful process discovery algorithms, which is largely rooted in the domains of computer science and business process management. A parallel is drawn between the process discovery requirements and audit concerns such as the need for efficient audit procedures that are capable of consistently and transparently delivering sufficient, appropriate evidence of the presence (or absence) of material misstatements in a company's financial statements. A final set of process discovery requirements for audit practice is established, along with suitable approaches for their appraisal. The practical use of the requirements for the evaluation of current and forthcoming process discovery algorithms on their applicability in an audit context is demonstrated for four state-of-the-art process discovery algorithms.

Keywords: process mining, process discovery, evaluation measures, auditing

1. Introduction

The internal and external audit constitute key instruments in the analysis and validation of business processes and the information disclosed in the financial statements based on these processes. During the audit, insight is gathered into a company's processes, the ways in which actual process executions diverge from the expected or desired execution methods, the risks to the financial statements arising from these deviations and the effectiveness of the internal controls installed to counter these risks (Flores & Riquenes, 2020; Werner, 2017). However, audit procedures traditionally tend to transpire on a 'what you see is what you get' (WYSIWYG) basis. Information is provided by the auditee; in the form of registrations in the information systems, interviews, walkthroughs, or additional business documentation (Jans et al., 2010). The auditor will typically perform IPE ('Information Produced/Provided by the Entity') testing on this information, e.g. by examining the programmes, queries and parameters that generated the provided reports (Wiese, 2019). These tests nevertheless do not grant a solid guarantee against (potentially deliberately) incorrectly entered, altered, deleted or outdated data. Moreover, audit procedures generally rely on sampling, since the substantial cost in both time and financial means of analysing the extensive number of transactions of a company may preclude an integral examination (Jans et al., 2010; Werner, 2017). In light of these aforementioned limitations, for instance Flores and Riquenes (2020) and Jans et al. (2013, 2014) recognise the potential of introducing 'process mining' in an audit environment and herald the dawn of 'Audit 2.0', a profound reform of audit practice to benefit from process mining techniques and tools.

The domain of process mining is devoted to furthering businesses' understanding of their processes and supports both process monitoring and process improvement. This is achieved by means of an array of techniques and tools that serve to derive knowledge on processes from automatic registrations in the supporting information systems (so-called 'event logs') (De Weerdt et al., 2012; Jans et al., 2013; W. van der Aalst, 2016b). This study centres on process discovery, the subdomain of process mining concerned with reconstructing both the history of individual transactions and prevalent process behaviour from an event log (Jans & Hosseinpour, 2019). The merits of applying process discovery techniques in audit practice are manifold. Firstly, event logs are ubiquitous and rich in data. The majority of middle to large companies avail themselves of (process-aware) information systems, whose event logs usually feature timestamps, data on the system users and more (Jans et al., 2010, 2014). Secondly, process mining techniques allow the auditor to forgo sampling. In theory, all information system transactions could be taken into account, which would prevent the oversight of 'outliers', process executions that occur infrequently and could therefore fail to be a part of an audited sample. Outliers establish important audit information, as they could signify a defect in a company's internal control (Jans et al., 2014; Werner, 2017). Thirdly, process discovery surpasses the WYSIWYG basis. Assuming appropriate IT governance, contextual 'meta-data' are automatically recorded alongside the input data provided by the auditee. A more verifiable and reliable view of the process under consideration can therefore be obtained (Chiu & Jans, 2019; Jans et al., 2010, 2013). Lastly, process discovery can (partially) automate audit procedures. This permits auditors to dedicate their time to more fundamental, complex or risky tasks that require professional judgment (De Bonhome et al., 2018; Werner, 2017).

Notwithstanding their potential in audit practice, process discovery algorithms are inherently rooted in the domains of computer science and business process management (Jans & Hosseinpour, 2019). It could therefore be the case that some of their emphases, assumptions and evaluation criteria are at odds with audit concerns. By way of illustration, process discovery algorithms generally relinquish the visualisation of infrequent process executions in favour of creating a straightforward, legible overview of a process' mainstream behaviour. As stated above, exceptional process instances could nonetheless be valuable in an audit context through the delineation of undesirable or fraudulent behaviour, and should not be disregarded during the analysis (Genga et al., 2018). It is therefore this research's objective to assess *the extent to which previously established process discovery requirements cover audit concerns*. In terms of research contributions, an in-depth review of audit literature is conducted to construct a list of requirements that process discovery algorithms need to comply with in order to facilitate efficient and effective audit practices. Extending the literature review to process mining and process discovery publications, a translation is made between extant audit and process discovery requirements. The study expounds on which audit requirements were previously recognised or met by process discovery literature, and how the requirements can be evaluated in practice. By way of demonstration, four state-of-the-art discovery algorithms will be appraised on the newfound process discovery requirements for audit practice, using both their introductory publications and models they mined from a real-life event log.

The remainder of the paper is organised as follows. Section 2 provides a brief introduction to the domain of process mining and the conceptual case for its application in an audit context. Section 3 elaborates on the study's set-up. Section 4 presents an overview of the currently recognised requirements for respectively audit practice and process discovery algorithms. Section 5 embodies the crux of the study and assesses the extent to which requirements that process discovery algorithms need to satisfy in order to be valuable in audit practice are currently acknowledged by process discovery research. This is achieved by drawing a parallel between the previously identified audit and process discovery requirements. Evaluation methods and metrics are proposed for all extant process discovery requirements that were shown to be valuable in an audit context. Section 6 comprises a demonstration of the study's findings, and appraises several state-of-the-art discovery algorithms on the newly proposed process discovery requirements for audit practice. The study closes on a discussion of the encountered constraints and the conclusions that are put forward.

2. Background

The following section introduces the concept of process mining (2.1) and the case for its application in audit practice (2.2).

2.1. Process mining

The domain of process mining, 'business process mining' in full, was established in the early 2000s, and is situated on the frontier of data mining and business process management. From a data mining viewpoint, process mining techniques and tools endeavour to extract knowledge from the vast volumes of data residing in companies' information systems (e.g. ERP, WFM and CRM systems) (De Weerdt et al., 2012; Jans & Hosseinpour, 2019). These data come in the form of event logs, also known as 'execution logs' or 'audit logs'. Event logs constitute a chronological overview of the activities that transpired during the execution of processes supported and controlled by information systems (De Weerdt et al., 2012; Jans et al., 2013). More specifically, an event log can be defined as a collection of 'traces', where each trace represents an ordered sequence of events associated with a particular process execution. For each event, or task execution, data are (automatically) recorded on the type of task performed (e.g. 'sign', 'pay'), the overarching case or process instance (e.g. a purchase order, an invoice), and the time or relative order of execution (Măruşter et al., 2006; W. van der Aalst, 2016b). Regularly, additional information is logged as well, such as the name of the system user ('resource') carrying out the task or the associated transaction value (Jans & Hosseinpour, 2019). Depending on how much information is stored, the deductions made on the grounds of the event log will be more accurate. On the flip-side, additional storage activities could slow down the information system, exert a negative influence on employee privacy and demand more disk space and event log assembly efforts (Chiu et al., 2019; Jans et al., 2013).

From a business process management viewpoint, process mining seeks to obtain insight into business processes. This insight can appertain to the organisation and those who participate in the process ('organisational perspective'), the overall process execution and its most frequent behavioural patterns ('process perspective'), or the individual process instances and their histories ('case perspective'). Altogether, the information obtained through process mining serves to clarify, monitor and improve a company's actual process behaviour, as opposed to its ideal or expected process executions (De Weerdt et al., 2012). Five process mining approaches can be discerned. Firstly, process discovery infers a process model or process description from the events recorded in an event log. Hitherto, process discovery has been the focal point of process mining research and various algorithms have been introduced to develop comprehensible process models which can reconstruct the behaviour recorded in an event log (De Weerdt et al., 2012; Jans et al., 2010, 2013; Jans & Hosseinpour, 2019). Process discovery is this study's main topic. Next, conformance checks evaluate the degree of correspondence between actual process executions, as they occur in practice, and ideal process executions, as they are expected or prescribed to take place. Any deviations can be detected and examined in more detail. Thirdly, throughput time, resource utilisation and other KPIs are computed and assessed in 'process performance analysis'. Fourthly, 'organisational mining' explores a company's organisational structure, task and role allocation, and interrelations between the employees and departments performing process steps. A last process mining approach is 'decision mining'. Here, the decision making within processes is investigated. For example, conditions for engaging on different process paths are derived or validated (Jans et al., 2013).

2.2. Process mining in an audit context

As already touched upon in the introduction, process mining applications are mainly situated in areas such as business process management, IT security and healthcare. Audit practice has as of yet mostly refrained from fully tapping into the techniques' potential (Jans et al., 2014). Nonetheless, when compared to traditional audit procedures, process mining approaches could succeed in uncovering more potential control issues that ought to be brought to the auditor's attention. This is because process mining expands the audit investigation from mere samples of the company's transactions to its full collection of information systems records. The discrepancy between analytical procedures and tests of detail is thereby eliminated, and the odds of detecting infrequent, abnormal behaviour are increased. Aside from user input on the executed activities, the information system records equally include meta-data, i.e. complementary, automatically logged data on for example the time of execution or the performer of an activity. Meta-data are currently limitedly availed of in an audit context, yet remain ubiquitous and valuable for retrieving information independently of the firm under investigation (Jans et al., 2013, 2014). So far, Jans et al. (2013, 2014) and Flores and Riquenes (2020) have already made a conceptual case for the implementation of the five previously explored process mining approaches (section 2.1) in audit practice.

By offering start-to-end views of a business' processes, process discovery presents a more efficient and comprehensive version of the traditional audit walkthrough procedure (Chiu et al., 2019; Jans et al., 2013). Moreover, process discovery promotes the analysis of real-life 'variants' of the processes of interest. Variants are unique execution patterns of a process or, expressed differently, process instances' distinctive histories. Unexpected or undesired variants could indicate an inefficiency or a failure of controls, that warrants further audit investigation depending on the variant's materiality (Chiu & Jans, 2019). In a similar way, conformance checks allow the auditor to compare observed process behaviour to either a predefined model expressing desired process behaviour or a set of business constraints. Both desirable and undesirable divergences will be located for further examination (Jans et al., 2013). Thirdly, process performance analysis can among other things be used for the examination of process instances' throughput times. Suspiciously protracted or short process instances, e.g. compared to average throughput times, call for further inspection to ensure that all process controls are effective (Chiu & Jans, 2019). Organisational process mining approaches could aid an auditor in the evaluation of a business' segregation of duty controls and the exposure of collusive fraud. For example, process mining techniques could verify whether crucial activities such as 'Create PO' and 'Sign PO' are effectuated by separate employees in all observed process instances (Jans et al., 2013, 2014). Moreover, social networks could disclose the intrinsic entanglements (e.g. handovers, invoicing, authorisations) between a selection of company stakeholders (e.g. employees, suppliers) (Jans et al., 2013). The last process mining approach, 'decision mining', could serve to derive or verify the conditions guiding a choice for a particular process path. Based on materiality, a deviation from standard practice could necessitate further audit investigations (Jans et al., 2010, 2013).

3. Research Design

As stated in the introduction, Audit 2.0 has the potential to pave the way for audit procedures of increased quality, rigidity and efficiency (Flores & Riquenes, 2020). However, to facilitate a successful integration of process mining techniques and tools in audit practice, current and forthcoming process discovery algorithms ought to suit audit requirements. The following research question therefore presents itself: *'To what extent do currently established requirements for process discovery algorithms cover audit concerns?'*.

An extensive literature review was conducted to address the proposed research question. Following the definitions, framework and guidelines of vom Brocke et al. (2015), the literature review was both systematic and iterative in nature. This means that the literature review systematically adhered to a chosen, explicit method, and iterated between searching and examining publications. Overall, the literature review consisted of two stages. In the first stage, requirements for effective and efficient audit procedures were identified by perusing the International Standards on Auditing (ISAs)[1], the financial audit standards issued by the International Federation of Accountants (IFAC). The objective was to comprehensively discern relevant audit requirements. Relevancy was gleaned from the frequency at which the requirements were mentioned, the urgency with which they were stressed and the extent to which they were elucidated.

In the second stage, contemporary process mining literature was examined in order to grasp the current perspective on process discovery requirements

[1] www.iaasb.org

and their metrics. An iterative keyword search was grounded in the UHasselt university library e-resources (i.e. UHasselt Discovery) and Google Scholar. The former provides access to 111 electronic databases, such as ACM Digital Library, IEEE Explore and Scopus. Search strings were generated progressively: first by using search terms directly related to the research question (e.g. 'process discovery quality'), and subsequently by expanding on plausible process discovery requirements deduced from previously inspected publications (e.g. 'process mining noise', 'process model understandability' and 'process mining precision'). With respect to search parameters, filters were applied on publication language, publication date, and publication type. English was the publication language of choice, since this language is understood by the author and covers a large section of process mining and process discovery literature. Publication dates were restricted to the period between 2005 and 2021, as process mining is a relatively recent and evolving field (De Weerdt et al., 2012). By focusing on the last sixteen years, a manageable set of publications is obtained, and precedence is given to requirements listed for contemporary discovery algorithms. A last set of parameters targeted the search at peer-reviewed journal articles and conference papers, of which the full text was available. This was implemented to foster the quality of the retrieved publications. Candidate publications obtained through the keyword search were ultimately selected for further examination based on the relevancy of their title and abstract to the research question. Approved publications were complemented with the results of backward citation searches, which were equally screened on recency and relevancy. This way, any significant publications that were overlooked during the original keyword search, could still be retrieved. From all publications, process discovery requirements were extracted based on the frequency at which they were mentioned, the urgency with which they were stressed and the extent to which they were elucidated.

The literature review was considered to be completed once additional ISAs and process mining publications no longer introduced any valuable new requirements. At that point, all identified process discovery and audit requirements were set side by side to formulate a response to the research question. Similar requirements were labelled as matches insofar as applicable. Conversely, audit requirements of which no related process discovery requirements could be identified, were recorded to be gaps in the process mining literature. From the previously collected literature, established evaluation methods and metrics were derived for the process discovery requirements that were deemed valuable in an audit context.

Finally, by way of demonstration, four state-of-the-art process discovery algorithms were evaluated on the requirements that ought to be satisfied by discovery algorithms in order to be valuable in an audit context. To this end, both their introductory publications and models mined from a real-life event log were taken into account, as further explained in section 6.1.

4. Identification of requirements

The following section provides an overview of previously acknowledged audit requirements (4.1) and requirements for valuable process discovery algorithms (4.2).

4.1. Audit requirements

Through the generation and evaluation of audit evidence, an auditor seeks to appraise the true and fair view of the financial statements and their compliance with the applicable financial reporting framework (ISA 700). Doing so, he ought to adhere to several requirements, which will be discussed in this section.

4.1.1. Ability to verify management assertions

In order to establish an opinion on the fairness of the financial statements and their compliance with the applicable financial reporting framework, an auditor should be able to identify material misstatements. A misstatement is a recorded financial statement item that deviates through error or fraud from the demands set by the financial reporting framework, e.g. about the magnitude or presentation of the item. Material misstatements are misstatements that, considered either individually or collectively, can be expected to influence the decision-making of a group of average users of the financial statements (ISA 320).

Misstatements can be uncovered through verification of the 'financial statement assertions', also known as 'management assertions'. These are five categories of implicit and explicit claims by a company's management regarding the financial reporting framework. The first category, 'completeness', demands that all transactions, assets, liabilities and equity interests of the company are included in the financial statements. Secondly, 'occurrence' requires that only transactions, assets, liabilities and equity interests that truly exist and belong to the company are taken into account. The third assertion, 'cut-off', indicates that all incorporated transactions should have occurred within the accounting period under consideration. 'Accuracy' specifies that all provided transaction data should be correct, and all assets, liabilities and equity interests should be recorded at a suitable value. Finally, 'classification', requires all transactions to be reported in the proper accounts and all assets, liabilities, and equity interests to be properly allocated (ISA 315).

4.1.2. Sufficient audit evidence

To formulate an opinion on the fairness of the financial statements and their compliance with the applicable financial reporting framework, the auditor needs to gather sufficient audit evidence of the presence (or absence) of material misstatements. He hereby takes into account all significant events preceding the publication of the audit report and all information attached to the financial statements and audit report that could affect the financial statements' credibility (ISA 560, 700, 720). 'Audit evidence' refers to the information collected from both executed audit procedures and other sources, that is cumulative in nature and serves to aid the auditor in formulating an opinion and constructing an audit report (ISA 500). By accumulating audit evidence, the auditor seeks to lower 'audit risk' to an acceptable level and obtain 'reasonable assurance' to state his opinion. Audit risk is the risk of an auditor not recognising the financial statements to be materially misstated. When stating an opinion, auditors can only strive towards reasonable and not absolute assurance due to inherent limitations of the audit, such as uncertainty about monetary value estimates and uncertainty about the integrity and interpretation of information provided by the auditee (ISA 200).

The quantity of audit evidence deemed 'sufficient' depends on several factors. Firstly, more evidence will probably be desired in case of a high perceived risk of material misstatement, be it due to a high likelihood or a large potential impact of an encountered misstatement. In a similar sense, indications of fraud require the auditor to increase his watchfulness, as they equally call into question the reliability of other evidence provided by the auditee and amplify the risk of concomitant additional fraud instances. Secondly, the higher the quality of the collected evidence (see also section 4.1.3), the less evidence might be requested. Contrarily, a high amount of audit evidence cannot entirely offset low evidence quality (ISA 500). Ultimately, it is up to the auditor to decide whether sufficient evidence has been collected, relying on his professional judgment of the nature of the evidence gathered, the expected costs and benefits of additionally collected evidence, and the time frame of the audit (ISA 200, 240).

As a general rule, auditors often fall back on sampling. This means that audit procedures are only conducted for a subset of the financial statement items, and the results are afterwards used to gauge the degree of misstatement of the entire population. In case of sampling, it is important to gather a representative subset of a suitable size, in order to minimise the risk of misapprehending the composition of the population ('sampling risk'). Depending on the context, the auditor can equally effectuate sampling by selecting specific items that for instance represent the largest monetary values, carry the largest risks, appear unusual or have features in common with previously identified instances of fraud. The results obtained through this type of sampling cannot be generalised to the item population, but selecting non-standard instances reduces the list of items to be tested to a reasonable size and focuses the auditor's attention and resources on the items that carry the largest risk of material misstatement (ISA 530) (Chiu & Jans, 2019).

In the absence of sufficient audit evidence, the auditor will be forced to either issue a qualified opinion or disclaim an opinion, depending on the estimated impact of potentially overlooked misstatements (ISA 330).

4.1.3. Appropriate audit evidence

The evidence gathered by the auditor in order to determine with reasonable assurance whether there exist any material misstatements in the financial statements, should not only be sufficient, but also appropriate. Evidence's 'appropriateness' is assessed through its relevance and reliability (ISA 200). In order to be relevant, the evidence needs to be logically connected to the financial statement item under consideration. For example, outstanding invoices and responses of suppliers to inquiries could play an important role in investigating a potential understatement of the accounts payable (ISA 500).

Audit evidence's reliability depends on the source from and conditions under which the evidence is collected. Potential sources are the company's collection of accounting records, past audits, employees, experts or other third parties. Generally, it is advisable to use various sources and types of evidence, and preference is given to independent sources external to the auditee, sources internal to the auditee that are regulated by effective, tested controls or evidence generated by the auditor himself through audit procedures (ISA 500). Written evidence is favoured over oral evidence, and is assumed to be authentic and unaltered in the absence of clear indications to the contrary. Depending on the source of a piece of evidence, additional analyses can be imposed to increase reliability. For example, to be able to adopt records of past audits, the auditor needs to ensure that no significant changes have transpired since their creation (ISA 500, 540). Reliability can be further improved by the availability of benchmarking data, such as industry data that the audit evidence can be compared to (ISA 520).

The desired level of appropriateness of audit evidence depends, among other things, on the perceived risk of material misstatement. The significance of this risk is influenced by both the likelihood and the potential impact of encountered misstatements (ISA 330).

4.1.4. Audit efficiency

While performing an audit, the auditor should guard the efficiency of the executed procedures. Since the relevance of information tends to diminish over time, a suitable balance needs to be maintained between the duration, cost, difficulty and reliability of the audit (ISA 200). Moreover, the procedures will be performed within a certain time frame and with limited human and financial resources. Adequate attention should therefore be paid to the planning of the audit, to ensure an appropriate assignment of resources to the most crucial components (ISA 300). Crucial components can be identified through an assessment of the risk of material misstatement of the various financial statement items. This risk depends on the probability of having a material misstatement for a particular type of financial statement item (the 'inherent risk') and the effectiveness of the business' preventative, detective and corrective controls (the 'control risk'). Procedures to determine the risk of material misstatement can include, among other things, a preliminary exploration of the business and its environment, inquiries of management and employees, observations, data analysis and inspections (ISA 200, 320, 330) (Jans et al., 2013; Putra, 2009). Taking into account the estimated risks of various transaction classes, account balances and disclosures, the auditor determines the nature, timing and scope of the audit procedures to be executed (ISA 300). However, it should be remembered that audit planning is iterative in nature. Throughout the audit, information about the auditee's internal and external environment should be duly updated and the audit strategy and plan should be revised when necessary. This way, the auditor can persist in ensuring optimal efficiency of the procedures executed during the audit investigation (ISA 315).

4.1.5. Consistency and transparency

The credibility and comprehensibility of an audit are generally enhanced by consistency in both evaluation approach and reporting style. In particular, congruent content and form allow a user to relatively quickly evaluate the audit's compliance with the applicable financial reporting standards (ISA 450, 500). Nevertheless, the investigated items, applied sampling, executed audit procedures, and timing and extent of the audit should preferably vary throughout time, with a view to remaining unpredictable for the auditee and lowering the odds of a successful fraud attempt (ISA 240).

A second important aspect is the transparency of the audit. The auditor should be able to provide substantial audit documentation explaining the executed procedures and justifying the final audit opinion to any unassociated, experienced auditor. More specifically, the auditor needs to elucidate the performed planning and risk analysis, the timing and extent of the audit, the tests executed and their performers, the (sometimes more subjective) evaluations made, the legal and regulatory framework, the evidence obtained and conclusions drawn, and any particular discrepancies or circumstances encountered. The generated documentation serves several purposes, such as guiding the planning process, facilitating quality reviews and establishing a paper trail and accountability (ISA 230).

4.1.6. Trustworthiness

In charging the auditor with the task of expressing an opinion on the fairness and compliance of the financial statements, the company's shareholders trust him to make a scrupulous judgement and to safeguard all data gathered while conducting the required procedures. Hence, aside from adhering to the relevant ISAs, the auditor should equally abide by the ethical requirements of the IFAC Code and all applicable (inter)national laws and regulations (ISA 200) (De Bonhome et al., 2020).

The International Federation of Accountants' Code of Ethics for Professional Accountants (IFAC Code) requires the auditor to exhibit integrity, objectivity, confidentiality, competence and professional behaviour. To this objective, the auditor should be independent of the auditee. The application of prior training, knowledge and experience in decision-making ('professional judgment') should be exerted during risk assessment, audit planning, evidence evaluation, and opinion formulation. In addition, professional skepticism should be demonstrated towards assumptions made during the audit, information provided by the auditee and third parties, contradictory audit evidence, and any circumstances potentially indicating misstatements and fraud (ISA 200).

Last but not least, the ISAs do not overrule (inter)national jurisdiction. The auditor should therefore make sure to continually operate in accordance with any applicable legal and regulatory obligations outside the ISAs (ISA 200, 700). For instance, auditors are responsible for safeguarding their clients' data. In general, statutory auditors are considered to be 'data controllers' under the General Data Protection Regulation (GDPR). When generating audit evidence, auditors often decide for themselves what data should be provided by the auditee, and to what purpose. Any of these data which are either stored in the European Union and in relation to an identifiable person or connected to an identifiable person who is located in the European Union, fall under the protection of the GDPR. This means that the auditor should proceed carefully when deciding how much data are to be processed and stored (i.e. a minimal amount), for what legally justified purpose, and how this is to be realised (i.e. securely, confidentially, transparently to the data subject, for a minimal amount of time and in consideration of any potential risks to the data subject) (De Bonhome et al., 2020; "GDPR one year on: its impact on auditors and accountants?", 2019; "GDPR", 2018; Gillin & Spoor, 2018).

4.2. Process discovery requirements

Over the last two decades, process mining literature has predominantly focused on the subdomain of process discovery. Under influence of this scrutiny, and originally drawing from a computer science and business process management background, numerous heterogeneous process discovery algorithms have been introduced to the field. On top of that, a diverse set of criteria has been established to assess the algorithms' usability and proficiency in expressing the behaviour of a process based on the records in an event log (Augusto, Conforti, Dumas, La Rosa, Maggi et al., 2019; De Weerdt et al., 2012; Werner, 2017). From these criteria, a number of requirements for valuable process discovery algorithms can be derived. In what follows, these requirements will be further explored.

4.2.1. Fitness, noise and the extent of event log behaviour visualised

The 'fitness' measure indicates how much of the behaviour observed in the event log is depicted in the process model. A fitness of 100% implies that all transactions in the event log follow a path integrally portrayed by the process model (Jans & Hosseinpour, 2019).

The overall fraction of event log behaviour visualised is, among other things, influenced by the process discovery algorithm's approach to noise. In process mining literature, noise is alternately defined in the spirit of 'interferences', 'deviations' or both, as interferences and deviations are generally indistinguishable when solely relying on the event log (W. van der Aalst, 2016b). When noise is defined along the lines of interferences, it is the result of incorrect logging and includes all things ranging from executed operations that failed to be recorded, to failed operations that were nonetheless registered, to random damage to the entries in the event log and incorrect timestamps (Mitsyuk & Shugurov, 2016). According to this interpretation, the approach to noise is more of a remedy to event log quality problems than a direct and voluntary effort to regulate the fraction of process behaviour visualised. Therefore, section 4.2.3 will further elaborate on interferential noise. When noise is defined in consonance with deviations, it relates to authentic, infrequent process executions. These infrequent process executions cover three categories: 'exceptions', 'anomalies', and 'fraud'. Exceptions are valid, acceptable deviations, which enable a business to respond more flexibly, efficiently or effectively to operational needs. An example is the forsaking of an additional inspection activity in case of a well-known supplier. Anomalies, also known as operational errors, are undesirable deviations caused by inadvertent deficiencies. Depending on legislation and company policies, anomalies amount to either mere inefficiencies or violations. An example of an anomaly is an employee accidentally recording the same invoice twice. Lastly, fraud constitutes behaviour that is undesirable, yet carried out deliberately and often covertly. An example is an employee paying an excessive price for ordered goods, and subsequently sharing the overpaid amount with the supplier (Baader & Krcmar, 2018; Depaire et al., 2013; W. van der Aalst, 2016b). The distinction between these three categories of deviant behaviour is mainly acknowledged in the context of the 'conformance checking' approach to process mining (discussed in section 2). Here, process instance patterns, transaction characteristics and rules are specified and learnt with an eye to distinguishing the different deviation types (e.g. Baader and Krcmar (2018), Jans and Hosseinpour (2019)). Within process discovery, chiefly the umbrella term 'noise' is used, without any further cataloguing.

Since event logs are often extensive and unstructured, discovery algorithms tend to lean towards process models featuring only the most mainstream, representative behaviour. Considering process mining's roots in process management, implementing changes to the most prevalent behaviour will frequently have a relatively large impact on the business (Bose et al., 2011; De Weerdt et al., 2012; W. van der Aalst, 2016b). Furthermore, solely focusing on the most common behavioural patterns generally results in more compact depictions, effectively lowering the models' visual complexity and the time and effort required of a user to understand them (as explained in section 4.2.4) (Cheng & Kumar, 2015). However, useful information could be lost by neglecting deviant behaviour. Deviant process instances could indicate new, advantageous paths to be considered for addition to the model representing all desired or acceptable process behaviour (in case of exceptions) or deficiencies in internal controls that needs to be remedied (in case of anomalies and fraud) (Genga et al., 2018).

4.2.2. Precision, generalisation and the completeness assumption

While generating a process model, process discovery algorithms make explicit and implicit assumptions about the underlying event data. One of these assumptions relates to the completeness of the event log with respect to the spectrum of all feasible process behaviour. Algorithms with strong completeness assumptions presume the event log to be comprehensive and to encompass at least one example of every possible type of process behaviour. Weaker completeness assumptions concede that a process may have a myriad of feasible execution methods, that take place with disparate probabilities. Hence, the event log will only feature a limited sample of all possible process behaviour (W. van der Aalst, 2016b).

Related to the discovery algorithm's outlook on event log completeness is its approach to behaviour that is not described in the event log. This approach is appraised by the 'precision' and 'generalisation' measures. 'Precision' reflects the degree to which the behaviour suggested by a process model is substantiated by evidence in the event log. A precision of 100% indicates that the behaviour allowed by the process model is strictly confined to behaviour observed in the log. In pursuing high precision values, discovery algorithms adopt stronger completeness assumptions and seek to avoid so-called 'underfitting' models. Underfitting models visualise arbitrary behaviour unrelated to the process under consideration (Jans & Hosseinpour, 2019; W. van der Aalst, 2016b). 'Generalisation', on the other hand, refers to the ability of a process model to relay behaviour that is implied, but not accounted for by the event log. Aiming for high generalisation values indicates a weaker completeness assumption and a wish to counter ‘overfitting’ models. Overfitting models solely allow the behaviour present in a single event log and are unlikely to capture new, at this point still unobserved instances of the process (W. van der Aalst, 2016b).

4.2.3. Ability to deal with log quality problems

For a large part, process discovery algorithms trust in the quality of the event logs they adopt as input. However, inadvertent errors could still make an appearance, in the form of faulty recording sensors, random damage to the entries in the event log, flawed integration of heterogeneous data sources, batch recordings of many user actions with identical timestamps, and more (Mitsyuk & Shugurov, 2016; Nguyen et al., 2019). Considering that the adage ‘garbage in, garbage out’ is equally applicable to process mining, hinging on low-quality, unrepresentative data could have disastrous effects on the efficacy of process mining techniques and tools (Wynn & Sadiq, 2019). Hence, to obtain useful, comprehensive and trustworthy process models, the event log should not only be complete (as explained in section 4.2.2), but also free of interferences (as mentioned in section 4.2.1). Due to particular characteristics of an event log, such as a multi-layered structure and dependencies between events as a consequence of cause-and-effect relations and resource availability, traditional data mining operations are not always suitable for ensuring event log quality (Nguyen et al., 2019; Suriadi et al., 2017). Therefore, process mining literature offers dedicated operations for improving event log quality while minimising information loss (Wynn & Sadiq, 2019).

Event log quality can be enhanced both during and after data acquisition. During data recording, attention should be paid that all cases and events featured in the event log have assuredly taken place in reality, unwarranted duplicates are avoided, and any reported attributes (e.g. timestamps, system user, transaction value) are accurate. Moreover, all attributes should be recorded with adequate granularity: sufficiently precise for the purpose of the analysis, but not too fine-grained for practical use. Lastly, any recordings in the information system should be consistent. This indicates that without any explicit notice of the opposite, activity labels and user names should remain stable and should continue to refer to the same unique objects and people (Suriadi et al., 2017; W. van der Aalst, 2016b). Data quality issues that failed to be prevented during acquisition, can be remedied through data cleaning and reconstruction procedures. Data cleaning is the process of identifying and filtering out queer, anomalous values in the input data (Nguyen et al., 2019; Wynn & Sadiq, 2019). These irregularities can be removed based on for example trace frequency, clustering, classifier rules or expert information (examples in Cheng and Kumar (2015), Li et al. (2018), Măruşter et al. (2006)). However, as explained in section 4.2.1, care should be taken when erasing logging errors, as these are often indistinguishable from authentic, infrequent behaviour (W. van der Aalst, 2016b). Data reconstruction entails the replacement of missing values in the event log. Missing values are cases, events and attributes that were lost as a consequence of recording failures or data cleaning procedures. Approaches such as the autoencoder method endeavour to recreate the missing input, even without any prior knowledge of the process control flow (Nguyen et al., 2019).

Up to this point, data acquisition, cleaning and reconstruction practices have mostly been considered preprocessing activities. Aside from a number of algorithms that include the removal of noisy instances (e.g. Günther and van der Aalst (2007), Mitsyuk and Shugurov (2016)), it appears that process discovery algorithms seldomly comprise steps that are dedicated to the evaluation or enhancement of event log quality.

4.2.4. Ability to simplify process visualisation

Simplicity measures evaluate the structural complexity of a discovered process model on various elements, such as model size or the amount of branching (Mitsyuk & Shugurov, 2016). As a rule, preference should always be given to the simplest model capable of representing the desired behaviour. This is because structural complexity, along with several other factors such as modelling approach, visual layout, domain familiarity and modelling expertise, relates to the mostly still equivocal measure of 'understandability' or 'comprehensibility'. Understandability covers the time and effort required of a general user to read, correctly interpret and use a process model. In addition, simpler models have a lower probability of containing errors and deficiencies (Dikici et al., 2018; W. van der Aalst, 2016b).

Many factors could affect the simplicity of a process model. One major source of structural complexity is the amount of (disparate) behaviour that the model is urged to represent. Blending many process executions of varying correspondence as if they were equally disparate could result in unduly large, spaghetti-like models, which are difficult to comprehend for users, fail to stress particular relationships between activities and increase the likelihood of incorporating arbitrary behaviour. A current approach to tackling this problem is to cluster process instances based on edit distance. Process discovery is performed separately on clusters that contain instances with high mutual similarity and high dissimilarity to instances in other clusters. When expressing similar execution behaviour, process variants can be combined more easily and coherently and model complexity will be reduced (Bose & van der Aalst, 2009a).

Another factor that could impact process model simplicity, is a discovery algorithm's approach to the identification, expression and abstraction of recurrent local behavioural patterns. Local behavioural patterns are subsets of activities that are frequently executed together within process instances. For example, a pattern might be a number of fine-grained activities related to image processing within a medical process, a request for further information following the preacceptance of a loan application, or the receipt of an invoice following a partial delivery. Behavioural patterns could be spread and repeated across the span of a process instance, could be enveloped by other, unrelated events, could be involved in more complex relationships such as concurrency and choice, could be varied upon and could be built from smaller patterns. Since process discovery algorithms tend to focus on the creation of end-to-end models, local behavioural patterns might sometimes be neglected. This disregard could once again trigger excessively detailed and complex process models that fail to emphasise relevant relations between activities (Bose & van der Aalst, 2009b; Genga et al., 2018; M. Leemans & van der Aalst, 2015; Tax et al., 2016). Currently, techniques within local process model mining and pattern mining (e.g. episode mining) are used to further the understanding of local relationships between activities (Bose & van der Aalst, 2009b).

In conclusion, many different aspects of a discovery algorithm's implementation could influence its performance on measures of structural process model simplicity such as model size and density. For example, it might be worthwhile for discovery algorithms to apply a form of clustering or pattern mining to improve the capture of similarities between process variants and local behavioural patterns. The presence of these types of techniques could therefore signal a discovery algorithm's heightened attention to model simplicity.

4.2.5. Practicality

Another important evaluation criterion for process discovery algorithms is their practical usability. To promote the application of process mining techniques in practice, their algorithms' demands in computation time and effort need to remain relatively low. In case of real-life processes, this entails fluent dealings with often vast and complex event logs (Augusto, Conforti, Dumas, La Rosa, Maggi et al., 2019). Although data storage and analysis are becoming ever the more cheap, attention should equally be paid to the required extent of logging. The latter comprises a trade-off between the value of additionally recorded meta-data and their effect on system speed, employee privacy and the effort of extracting and assembling data from various locations in the IT systems (Jans et al., 2010, 2013).

5. Audit requirements for process discovery algorithms

As discussed in the previous section, process discovery algorithms can be evaluated and improved on a plethora of requirements. Since process mining originated from a computer science and business process management background, section 5.1 seeks to gauge the extent to which these established process discovery requirements uphold audit concerns (Jans & Hosseinpour, 2019). Starting from this translation between process discovery and audit requirements, section 5.2 finalises a list of process discovery requirements that are valuable in an audit context, and explains how these requirements can be evaluated.

5.1. Translating audit concerns to process discovery requirements

The translation between the audit and process discovery requirements can be found in table 1. Over the course of the audit procedures, the auditor needs to gather evidence of the presence or absence of material misstatements in the financial statements. Evidence will therefore be dedicated to the confirmation or repudiation of five categories of *management assertions* (listed in section 4.1.1) (ISA 315). Process models discovered from the event logs created by an enterprise information system could aid in this endeavour. Assuming proper IT governance, the event logs comprise records of all transactions relevant to the company. Hence, the 'completeness' and 'occurrence' assertions could be verified using process models mined by process discovery algorithms sporting high fitness values, high precision values, low generalisation values and a correct approach to noisy, infrequent behaviour. These models contain most (or all) of the observed event log behaviour, and (virtually) no unsubstantiated behaviour. Transactions included in the models, which are absent in the financial statements, could be noted as potential

Process discovery requirements	Audit requirements: Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Fitness	+	+					
Approach to noise	+	+		+			
Precision	+		+	+			
Generalisation	-		-	-			
Ability to deal with log quality problems			+				
Simplicity			+	+		+	
Practicality				+			

Table 1: Translation between audit and process discovery requirements

breaches of the 'completeness' assertion. Contrarily, transactions incorporated in the financial statements, that failed to be part of the mined process models, suggest violations of the 'occurrence' assertion. The potential transgressions will nevertheless still need to be subjected to the auditor's professional judgement for corroboration. For the other assertions and the completeness and occurrence of assets, liabilities and equity interests, no direct connection to the process discovery requirements could be found. The 'accuracy' of the information included in the financial statements about the company's transactions (e.g. transaction values) could nonetheless be verified by matching the statements to the automatic recordings in the event log. Similarly, the verification of the 'cut-off' assertion might be facilitated for a specific type of process (e.g. purchase process), by exploring the process models built from all the transactions for which a specific activity (e.g. 'goods received') took place during the accounting period.

Next, the auditor should be able to provide *sufficient evidence* of the presence (or absence) of material misstatements (ISA 500). This can be linked to the fitness requirement of process discovery algorithms, as the visualisation of most (or all) of the transaction executions in the event log helps to ensure that an adequate amount of evidence is gathered. Once again, infrequent, irregular behaviour should equally be included in the discovered models, as these noisy instances could provide evidence of ineffective internal controls or fraudulent transactions. A final remark concerning the sufficiency requirement, is that the application of process mining could elevate audit evidence quality. Less evidence will therefore be required when formulating a grounded opinion on the fairness of the financial statements (ISA 500). However, since process discovery in theory allows the analysis of all transactions recorded in the information system's event log, sampling would no longer be required and the amount of evidence provided would be increased either way (Jans et al., 2013, 2014; Werner, 2017).

The *appropriateness of audit evidence* refers to its reliability and relevance to the financial statement item under inspection (ISA 500). None of the process discovery requirements has a direct impact on the reliability of the source and collection method of the gathered evidence. It can only be remarked that process discovery's automatically generated evidence could be assumed to be highly reliable in case of effective IT governance and high-quality event logs. Next, in view of the evidence's relevancy, it would be advisable to continually pursue high precision values and low generalisation values. This way, any process executions in the discovered models which are flagged to be incorrect, incomplete or unauthorised will assuredly have taken place in reality and will therefore have a greater probability of constituting relevant evidence to the investigation (Werner, 2017). Moreover, local behavioural patterns in a transaction's life cycle could preliminarily hint at the transaction's evidential relevance. This is because behavioural patterns could represent a specific type of process behaviour and particularities in a transaction's history that could indicate the presence of (material) misstatements.

Considering a backdrop of limited time and resources, auditors need not only be able to verify management assertions, they also need to be able to do so *efficiently* (ISA 300). To this end, process discovery algorithms should aim for high precision values and low generalisation values. This will ensure that all identified incorrect, incomplete and unauthorised process executions in the mined process models have in fact taken place in reality. Hence, no means will be wasted on investigating troubling process executions that are feasible, but as of yet unobserved (Werner, 2017). Secondly, an auditor could achieve greater efficiency levels by focusing on transactions with a higher risk of material misstatement. It is therefore worthwhile to include infrequent, noisy behaviour in the discovered process models, as irregular transactions could expose ineffectiveness of internal controls and (intentional) transgressions (Depaire et al., 2013; Jans et al., 2014; Werner, 2017). In a similar spirit, if transaction values are incorporated in the event log, the discovery algorithm could concentrate on process instances with a significant financial impact. The latter consideration is currently not integrated in any process discovery requirement, but might be interesting for future reference. Thirdly, the application of process discovery techniques in audit practice can only be efficient when discovered models are not excessively complex. The simplicity requirement, which states that the most simple and comprehensible visualisation of the behaviour in an event log should be obtained, is therefore essential (W. van der Aalst, 2016b). In addition, audit efficiency will be increased when running the process discovery algorithm demands only little computation time and effort. This will be even more important when new information about the auditee's internal or external environment is obtained sporadically over the course of the audit, urging the discovery algorithm to be run multiple times (ISA 315).

Consistency in audit evaluation approach and reporting style has no direct connection to the process discovery requirements. However, process discovery can contribute to consistency in form and content by building upon event logs of comparable quality, by exploring corresponding perspectives (organisational, process and case perspectives; see section 2), and by applying similar process discovery techniques, tools and parameters. Depending on the process discovery tool used, the executed analyses might even be scripted, saved and replicated. This is also related to the *transparency* requirement, which entails that an auditor should be able to explain and justify the procedures followed and evaluations made over the course of the audit (ISA 230). Furthermore, drawing upon the fact that discovered process models can be used as audit evidence and documentation, it is also beneficial to strive for comprehensible process models (i.e. high simplicity values). At any rate, the use of process mining in general may already lead to increased traceability of the executed audit procedures (De Bonhome et al., 2020).

The last audit requirement, *trustworthiness*, indicates that auditors should abide by the ethical requirements of the IFAC Code and all applicable (inter)-national laws and regulations such as the GDPR (ISA 200, 700). None of the process discovery requirements directly affect this audit requirement.

5.2. Metrics for process discovery requirements in audit practice

From the previous subsection's exercise, eight process discovery requirements for audit practice can be derived. Through consideration of these requirements, auditors can evaluate process discovery algorithms on their value in an audit context.

Firstly, process discovery algorithms should pursue high *fitness* and *precision* values (i.e. close to one). High fitness values will allow the auditor to assemble sufficient evidence of the presence (or absence) of misstatements in the financial statements. Moreover, he could avail himself of the discovered models when verifying that the financial statements are complete and all relevant items have been recorded. High precision values will increase the audit's relevance and efficiency, and facilitate the verification of the 'occurrence' assertion (Werner, 2017). Throughout the years, an assortment of metrics have been introduced for the evaluation of discovery algorithms' fitness and precision. A selection of these metrics, as recommended by comparative studies such as Janssenswillen et al. (2017) and Syring et al. (2019), are listed in Appendix 1.

Next, the discovery algorithm should enable the auditor to *regulate the amount of noise* included in the mined model. Allowing for little infrequent behaviour provides the auditor with an intelligible view of a process' most mainstream behaviour. This perspective is mostly instrumental in the audit planning phase. Contrarily, the examination of infrequent process behaviour increases audit procedures' efficiency, assists the auditor in verifying the financial statements' completeness, and aids in the collection of sufficient audit evidence. In the assembled literature, no metrics were found that had explicitly been designed for the purpose of measuring the amount of infrequent behaviour included by a discovery algorithm. However, a selection of preprocessing and discovery algorithms (e.g. Li et al. (2018) and Măruşter et al. (2006)) developed their own approaches for defining and quantifying noise with the aim of filtering out both infrequent behaviour and interferences. These approaches could be used as a source of inspiration for new (informal) noise metrics. Alternatively, the auditor could simply examine discovery algorithms' publications, with a focus on the applied definitions of noise and any implemented (adjustable) filters on infrequent behaviour.

A fourth requirement dictates that discovery algorithms should refrain from *generalising the observed behaviour* to process paths that are probable, but as of yet unconfirmed by the event log. This is to ensure both the audit's efficiency and the generated evidence's relevance to events that took place in the company. At the time of writing, only a limited number of generalisation metrics have been introduced to the field (e.g. alignment based probability (W. van der Aalst et al., 2012), behavioural generalisation (van den Broucke et al., 2014), and AVATAR (Theis & Darabi, 2020)). Moreover, these metrics regularly fail to be in agreement when evaluating discovered process models, which indicates that there is currently little consensus on how to quantify discovery algorithms' generalisation capabilities (Janssenswillen et al., 2017). Hence, the auditor might prefer to simply penalise any mentions of generalisation efforts in discovery algorithms' publications.

Fifthly, to ensure the relevance and accuracy of gathered audit evidence, process discovery algorithms should have a *suitable approach to event logs of limited quality*. Any references in their publications to measures taken for the identification and treatment of erroneous or missing values in an event log should be encouraged.

In the sixth place, process discovery algorithms should visualise the desired behaviour with maximum model *simplicity*, to foster the quality, efficiency and transparency of the applied audit procedures. As demonstrated in Appendix 1, there exists a multitude of simplicity metrics, each focusing on a limited aspect of structural model complexity. Hence, an overall evaluation of process model simplicity generally demands the implementation of a well-thought-out set of metrics befitting the context, the demands set by the analyst (i.e. the auditor) and his own interpretation of sources of process model complexity. The composition of this set is further complicated by the fact that for many simplicity measures, theoretical or practical validations have failed to take place. Moreover, some of the metrics are biased towards particular modelling languages or struggle with the evaluation of unstructured, real-life process models (De Weerdt et al., 2012; Polančič & Cegnar, 2017).

Efficient audit procedures furthermore expect any implemented process discovery algorithms to be *practical* in run time and computational effort. This can be ascertained by keeping track of the average time and working memory required for a discovery algorithm's computations (Janssenswillen et al., 2017).

Lastly, a new requirement identified in the previous subsection encouraged process discovery algorithms to differentiate process instances based on their transaction values. As a transaction's monetary value influences its *materiality*, future discovery algorithms could further an audit's efficiency by allowing the visualisation of event log behaviour to be guided by a user-defined filter on transaction value. As of yet, 'transaction value' is only an optional attribute in information systems' event logs, and process discovery algorithms pay relatively little attention to it.

6. Demonstration

In the following section, a demonstration will be given of how process discovery algorithms can be evaluated on their efficiency and effectiveness in an audit context. More specifically, four state-of-the-art discovery algorithms will be assessed on the proposed process discovery requirements for audit practice. Section 6.1 provides details on the demonstration's setup, whereas section 6.2 presents the obtained results.

6.1. Demonstration setup

This section elaborates on the demonstration's setup: the process discovery algorithms that were evaluated (section 6.1.1), the requirements that were assessed and the metrics that were applied (section 6.1.2), the event log that was used (section 6.1.3), and the visualisation that was created to present the demonstration's results (section 6.1.4).

6.1.1. Evaluated process discovery algorithms

At the outset of the demonstration, a collection of state-of-the-art process discovery algorithms was chosen for evaluation on the process discovery requirements for audit practice. Candidate algorithms were assembled by exploring process discovery publications of the last ten years for newly proposed algorithms. The subsequent election of algorithms was based on the recency of their proposal, their partaking in previous comparative experiments (e.g. Augusto et al. (2018), Janssenswillen et al. (2017)), and the availability of an implementation in ProM or on the Apromore platform. Four discovery algorithms were retained: the Inductive Miner, the Evolutionary Tree Miner (Buijs et al., 2014), the Split Miner (Augusto, Conforti, Dumas, La Rosa & Polyvyanyy, 2019), and the ILP miner (van Zelst et al., 2018). Of the Inductive Miner, both its standard version (S. J. J. Leemans et al., 2013) and its extension for dealing with infrequent behaviour (S. J. J. Leemans et al., 2014) were evaluated.

6.1.2. Evaluated requirements

The four selected process discovery algorithms were evaluated on five of the eight process discovery requirements for audit practice. A set of evaluation methods and metrics was already proposed in section 5.2. These recommendations are summarised in Appendix 1, and underlined according to their application in the demonstration. Ergo, each discovery algorithm's introductory publication was explored for information on the algorithm's approach to noise, measures taken to identify or rectify event log quality issues and generalisation efforts. The fitness and precision requirements were assessed by applying the negative event recall (Goedertier et al., 2009), alignment-based fitness (W. van der Aalst et al., 2012), one-align precision (Adriansyah et al., 2015) and alignment-based precision (W. van der Aalst et al., 2012) metrics on process models mined from a real-life event log. The event log is introduced in more detail in section 6.1.3. The applied metrics did not only perform well in comparative studies, but were equally chosen because of their standard implementation in either ProM (version 6.10) or CoBeFra (version of 2018-04-03). The plug-ins used are listed in Appendix 2. Notwithstanding its implementation in CoBeFra and positive rating in Syring et al. (2019), soundness was left out of the demonstration due to its inability to deal with looping behaviour.

The discovery algorithms were not (fully) evaluated on their simplicity, practicality and approach to transactions' materiality. As discussed in section 5.2, there exists an abundance of simplicity metrics, which cover different aspects of process model complexity (Polančič & Cegnar, 2017). Lacking guidance as to what sources of structural complexity chiefly impact a model's comprehensibility in face of an audience of auditors, the decision was made to confine the current evaluation to an inspection of the discovery algorithms' designated modelling languages. It should however be kept in mind that initially mined process models can still be converted to another modelling language. This is nonetheless at the risk of introducing errors or counterproductively increasing the model's structural complexity (Mitsyuk & Shugurov, 2016; W. van der Aalst, 2016a). Future research could seek to uncover what sources of structural complexity are cardinal in an audit context, and what simplicity metrics should therefore best be applied.

The practicality of the process discovery algorithm in computation time and effort was equally left out of the evaluation due to the fact that the demonstration's setting was largely artificial, involving only a sample of an event log and discovery algorithms with default settings. It is nonetheless important to remark that several publications agree that the Evolutionary Tree Miner generally exhibits relatively long execution times (Augusto, Conforti, Dumas, La Rosa, Maggi et al., 2019; Augusto, Conforti, Dumas, La Rosa & Polyvyanyy, 2019).

Lastly, the newly introduced requirement to differentiate process instances based on their transaction values, was not incorporated in the demonstration. As this requirement has yet to be put forward in process mining literature, none of the evaluated process discovery algorithms satisfied it.

6.1.3. Applied data

The event log used in the demonstration was originally presented in Chiu and Jans (2019) and Jans et al. (2014). It concerns a multinational European bank's procure-to-pay process, and contains the behaviour of 26,185 order lines of purchase orders that resulted in an invoice over the course of January 2007 (181,845 events, 7 types of activities). For reasons of efficiency, this study only considered a sample of 15,069 purchase order lines (103,720 events, 7 types of activities), by solely factoring in those order lines whose process execution started in the first half of January 2007. The plug-ins used to apply the process discovery algorithms on the event log, are listed in Appendix 2. Default parameter values were plied for all discovery algorithms, as parameter optimisation was considered to be outside the scope of the demonstration.

6.1.4. Visualisation of the results

The visualisation of the demonstration's results is based on the fusion of process discovery and audit requirements in table 1. Each figure provides the evaluation of a single discovery algorithm. More specifically, each row in the visualisation projects an algorithm's performance on a process discovery requirement that could affect the fulfilment of one or more audit requirements. A plus sign indicates the audit requirement's need for high scores on the related process discovery requirement, a minus sign signifies the opposite. Accordingly, a dark green score is considered to be advantageous, whereas a dark red score entails poor performance on the audit requirement. Greyed out cells imply that the process discovery requirement has no direct impact on the audit requirement. The final row presents aggregated results on all audit requirements. In their calculation, equal weights are assumed for all process discovery requirements. A follow-up study could seek to determine more accurate, relative weights to render the aggregation more nuanced.

6.2. Results

In what follows, the demonstration's main findings will be addressed. As explained in section 6.1.4, a graphical overview of the results can be found in figures 1 to 5.

6.2.1. Fitness and precision

The fitness requirement was satisfied for all evaluated process discovery algorithms. Average scores amounted to 0.8493 and 0.9628 for respectively negative event recall and alignment-based fitness, both nearing the maximum score of one. The best results were obtained for the Inductive Miner, while either the ILP Miner (in case of negative event recall) or the ETM (for alignment-based fitness) came in second place.

In similar fashion, all evaluated process discovery algorithms bar the standard Inductive Miner reported favourable results for the precision requirement. The Split Miner and ILP Miner obtained the maximum score for both metrics, and the ETM and Infrequent Inductive Miner returned promising results of on average 0.9222 and 0.8868 for one-align and alignment-based precision. Contrarily, the standard Inductive Miner only attained scores of 0.292 for one-align precision and 0 for alignment-based precision.

In conclusion, all evaluated discovery algorithms demonstrated a balanced approach to fitness and precision, performing satisfactorily on both fronts. Solely the standard Inductive Miner seemingly gives prominence to fitness at the expense of the resulting model's precision. This behaviour is confirmed by the algorithm's publication and partly toned down by its extension for curtailing infrequent behaviour (S. J. J. Leemans et al., 2013, 2014). Furthermore, it should be remarked that the ETM allows its users to rearrange the balance and focus its efforts on a quality dimension of choice (fitness, precision, generalisation or simplicity) through a number of preassigned weights. In addition, the ETM is a genetic algorithm, which iteratively generates candidate models in search of an optimal one. Its performance could therefore depend on the number of executed iterations (Buijs et al., 2014).

6.2.2. Approach to noise

Three of the evaluated discovery algorithms explicitly discussed their approach to noise. All three allow the analyst to regulate the amount of noise introduced in the model by means of one or more user-defined thresholds. More specifically, the Infrequent Inductive Miner allows users to define a parameter k, which determines the minimum frequency that direct and indirect ('eventual') relations between activities need to attain to be acknowledged in the model (i.e. k times the most prevalent outgoing relation of the first activity). Additional filters can on occasion be applied as well, for instance to remove infrequent activities (S. J. J. Leemans et al., 2013). The Split Miner applies two noise filtering parameters. The first removes any connection between two activities (a, b) that is observed significantly less frequently than the same connection in the opposite direction (b, a). The second, in similar fashion to the Infrequent Inductive Miner, establishes the threshold that the frequency of a connection between two activities needs to exceed for the relation to be included in the model (Augusto, Conforti, Dumas, La Rosa & Polyvyanyy, 2019). Lastly, the ILP Miner's noise filter equally focuses on infrequent connections between activities. It gradually reconstructs the body of constraints from which a process model is built, dropping parts of the constraints and their accompanying behaviour when their frequency of appearance is lower than a chosen threshold (i.e. $1 - \alpha$ times the most frequent constraint segment with the same prefix) (van Zelst et al., 2018).

6.2.3. Simplicity, generalisation and approach to log quality problems

Regarding the simplicity requirement, a maximum score was solely obtained by the Split Miner. The Split Miner generates a BPMN model, a type of model that is generally judged to be user-friendly, flexible, and rich in semantic (Aldin & de Cesare, 2009). In addition, the miner expressly endeavours to minimise branching complexity and frequently creates highly structured models (Augusto, Conforti, Dumas, La Rosa, Maggi et al., 2019; Augusto, Conforti, Dumas, La Rosa & Polyvyanyy, 2019). Both the Inductive Miner and ETM rate averagely on simplicity. They return process tree models, a type of model that tends to provide straightforward visual cues about a process' activities and their mutual relationships (Zhang, 2017). Lastly, the ILP Miner creates a Petri net model. Petri net models are often relied upon in process discovery due to their compatibility with metrics for process model evaluation. Nevertheless, the models are commonly considered to be little user-oriented (Aldin & de Cesare, 2009; Augusto, Conforti, Dumas, La Rosa, Maggi et al., 2019).

In the algorithms' publications, few references were made to deliberate generalisation efforts. This is in line with their high performance on precision, as reported in section 6.2.1. Moreover, none of the algorithms take any specific actions with regard to the evaluation or enhancement of event log quality. As mentioned in section 4.2.3, these efforts appear to be entrusted to preprocessing algorithms.

6.2.4. Overall evaluation

Combining the above-stated results across each audit requirement, optimal performances were realised by the Infrequent Inductive Miner, ILP Miner and Split Miner (see the final rows in figures 1 to 5). The Infrequent Inductive Miner and ILP Miner attained the best results for aiding in the verification of the management assertions and collecting sufficient evidence of the presence (or absence) of material misstatements. The Split Miner did not only achieve comparable results on these two requirements, but also returned the highest scores for the collection of appropriate audit evidence and the support of efficient and transparent audit procedures. Whereas the Split Miner's aggregated results were all greater than or equal to 0.75, the Infrequent Inductive Miner's and ILP Miner's aggregated scores fluctuated around this mark. The standard Inductive Miner and ETM consistently obtained aggregated scores below 0.75. It should nonetheless be remembered that the audit requirements' aggregated scores were obtained by assigning equal weights to all related process discovery requirements. Future research could indicate which process discovery requirements exercise the greatest influence on each audit requirement. Using fine-tuned weights, more nuanced conclusions may be obtained from the demonstration.

		Audit requirements						
		Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Process discovery requirements	Fitness	0.9274+	0.9274+					
	Approach to noise	0+	0+		0+			
	Precision	0.146+		0.146+	0.146+			
	Generalisation	0-		0-	0-			
	Ability to deal with log quality problems			0+				
	Simplicity			0.5+	0.5+		0.5+	
	Aggregated results	0.5184+	0.4637+	0.4115+	0.4115+		0.5+	

Figure 1: Evaluation Inductive Miner

		Audit requirements						
		Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Process discovery requirements	Fitness	0.9346+	0.9346+					
	Approach to noise	1+	1+		1+			
	Precision	0.9801+		0.9801+	0.9801+			
	Generalisation	0-		0-	0-			
	Ability to deal with log quality problems			0+				
	Simplicity			0.5+	0.5+		0.5+	
	Aggregated results	0.9787+	0.9673+	0.62+	0.87+		0.5+	

Figure 2: Evaluation Infrequent Inductive Miner

		Audit requirements						
		Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Process discovery requirements	Fitness	0.8862+	0.8862+					
	Approach to noise	1+	1+		1+			
	Precision	1+		1+	1+			
	Generalisation	0-		0-	0-			
	Ability to deal with log quality problems			0+				
	Simplicity			1+	1+		1+	
	Aggregated results	0.9716+	0.9431+	0.75+	1+		1+	

Figure 3: Evaluation Split Miner

Process discovery requirements	Audit requirements						
	Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Fitness	0.932+	0.932+					
Approach to noise	1+	1+		1+			
Precision	0.9995+		0.9995+	0.9995+			
Generalisation	0-		0-	0-			
Ability to deal with log quality problems			0+				
Simplicity			0+	0+		0+	
Aggregated results	0.9829+	0.966+	0.4999+	0.7499+		0+	

Figure 4: Evaluation ILP Miner

Process discovery requirements	Audit requirements						
	Ability to check management assertions	Sufficient evidence	Appropriate evidence	Efficiency	Consistency	Transparency	Trustworthiness
Fitness	0.8502+	0.8502+					
Approach to noise	0+	0+		0+			
Precision	0.8286+		0.8286+	0.8286+			
Generalisation	0-		0-	0-			
Ability to deal with log quality problems			0+				
Simplicity			0.5+	0.5+		0.5+	
Aggregated results	0.6697+	0.4251+	0.5822+	0.5822+		0.5+	

Figure 5: Evaluation Evolutionary Tree Miner

7. Discussion

This study has implications for the research fields of auditing and process discovery. By highlighting the added value of process discovery in an audit context, further contributions are made to the case for the application of process mining in audit practice as presented in Jans et al. (2013, 2014) and Flores and Riquenes (2020). Moreover, the proposed list of process discovery requirements for audit practice can support researchers in two ways. On the one hand, extant process discovery algorithms can be evaluated on their applicability in an audit context. On the other, the requirements can guide the development of forthcoming algorithms.

When interpreting the study's results, a number of limitations should nonetheless be kept in mind. First of all, the study solely focused on requirements for process discovery algorithms. For process discovery to be successfully introduced in an audit context, attention should equally be paid to the suitability of event logs and process discovery tools. Namely the event log creation process should be efficient, rigorous and transparent, and should culminate in a format that is practical for the subsequent analyses. Complicating factors might be that the required data are spread across various database tables, and that disparate data storage architectures inhibit automation of the creation process. Future work could therefore look into best practices for event log creation, and into satisfactory procedures for verifying whether event logs are both complete and reliable. Moreover, it should be determined what meta-data are required for evidence collection. For instance, some attributes are not mandatory for event log construction, but will nevertheless be essential in audit investigations (e.g. transaction values) (De Bonhome et al., 2020; Jans et al., 2013, 2014; Jans & Hosseinpour, 2019). Process discovery tools, on the other hand, should be user-friendly in face of an audience of auditors. Many auditors are still unfamiliar with process mining techniques, and should therefore be eased into their applications in order to minimise any additional workload (Jans et al., 2013). Furthermore, applied tools should be able to support all functionalities mentioned in section 5, such as a filter on the amount of noise and scripting of analyses.

Secondly, as of yet, not all proposed requirements can be assessed straightaway. Future research could therefore focus on developing formal, conclusive metrics for the approach to noise and generalisation requirements. Moreover, in order to facilitate the evaluation of the simplicity requirement, it should be established what sources of process model complexity are most significant to an auditor.

Thirdly, for the evaluation of discovery algorithms' performances on a specific audit requirement, the study assumed equal contributions of all related process discovery requirements. Ensuing research could seek to uncover their relative importance. In similar fashion, future work could estimate the relative prominence of each audit requirement in an overall assessment of a discovery algorithm's applicability. Establishing fine-tuned weights for all process discovery and audit requirements allows for more nuanced evaluations.

Lastly, the demonstration conducted in this study considered only a sample of a real-life event log, four process discovery algorithms with default settings and a selection of the proposed requirements. Hence, follow-up studies could seek to expand this demonstration, thoroughly evaluating an optimised discovery algorithm on more transactions and all requirements.

8. Conclusion

Conscious of process mining's potential in an audit context, this study contributed to the implementation of its flagship approach, process discovery. To fully rely on process discovery techniques and tools, auditors might seek reassurance that requirements for valuable process discovery algorithms reflect the standards and expectations of contemporary audit practice. Therefore, this study conducted a systematic review of the International Standards on Auditing (ISAs) and contemporary process mining literature, in order to derive extant audit and process discovery requirements. Both sets of requirements were placed in juxtaposition, and a final list was drawn up of requirements for process discovery algorithms in audit practice. The evaluation of these requirements was demonstrated for four state-of-the-art process discovery algorithms.

In order to be useful in an audit context, process discovery algorithms should aim for high fitness, precision and simplicity values, and low generalisation scores. Moreover, they should allow the auditor to regulate the amount of noise incorporated in the mined models and to filter visualisations based on transaction materiality. Lastly, the algorithms should be efficient in computation time and effort, and need to be able to detect and resolve event log quality issues.

Future research could look further into best practices for process discovery tools in an audit context. Furthermore, a number of requirements still lack formal or conclusive evaluation methods, and a more comprehensive corroboration of the proposed requirements should be obtained. Lastly, the relative importance of the different process discovery requirements for audit practice should be determined.

References

Adriansyah, A., Munoz-Gama, J., Carmona, J., van Dongen, B. & van der Aalst, W. (2015). Measuring precision of modeled behavior. *Information Systems and e-Business Management*, *13*(1), 37–67.

Aldin, L. & de Cesare, S. (2009). A comparative analysis of business process modelling techniques, In *UK Academy for Information Systems Conference Proceedings*.

Augusto, A., Armas Cervantes, A., Conforti, R., Dumas, M., La Rosa, M. & Reissner, D. (2018). Abstract and Compare: A Framework for Defining Precision Measures for Automated Process Discovery, In *International Conference in Business Process Management*, Sydney, Australia.

Augusto, A., Conforti, R., Dumas, M., La Rosa, M., Maggi, F. M., Marrella, A., Mecella, M. & Soo, A. (2019). Automated Discovery of Process Models from Event Logs: Review and Benchmark. *IEEE Transactions on Knowledge and Data Engineering*, *31*(4), 686–705.

Augusto, A., Conforti, R., Dumas, M., La Rosa, M. & Polyvyanyy, A. (2019). Split miner: Automated discovery of accurate and simple business process models from event logs. *Knowledge and Information Systems*, *59*(2), 251–284.

Baader, G. & Krcmar, H. (2018). Reducing false positives in fraud detection: Combining the red flag approach with process mining. *International Journal of Accounting Information Systems*, *31*, 1–16.

Bose, R. P. J. C. & van der Aalst, W. M. P. (2009a). Context Aware Trace Clustering: Towards Improving Process Mining Results, In *Proceedings of the 2009 SIAM International Conference on Data Mining*, Society for Industrial; Applied Mathematics.

Bose, R. P. J. C. & van der Aalst, W. M. P. (2009b). Abstractions in Process Mining: A Taxonomy of Patterns, In *International Conference on Business Process Management*, Berlin, Heidelberg, Springer.

Bose, R. P. J. C., van der Aalst, W. M. P., Žliobaitė, I. & Pechenizkiy, M. (2011). Handling Concept Drift in Process Mining, In *International Conference on Advanced Information Systems Engineering*, Berlin, Heidelberg, Springer.

Buijs, J. C. A. M., van Dongen, B. F. & van der Aalst, W. M. P. (2014). Quality Dimensions in Process Discovery: The Importance of Fitness, Precision, Generalization and Simplicity. *International Journal of Cooperative Information Systems*, *23*(1), 1440001.

Cheng, H.-J. & Kumar, A. (2015). Process mining on noisy logs — Can log sanitization help to improve performance? *Decision Support Systems*, *79*, 138–149.

Chiu, T., Brown-Liburd, H. & Vasarhelyi, M. (2019). Performing Tests of Internal Controls Using Process Mining: What Could Go Wrong? *The CPA Journal*, *89*(6), 54–57.

Chiu, T. & Jans, M. (2019). Process Mining of event logs: A case study evaluating internal control effectiveness. *Accounting Horizons*, *33*(3), 141–156.

De Bonhome, O., Jans, M., Gjymshana, E., Kroes, D., Marissen, M., Simpelaere, F., Trumpener, J. & Verachtert, S. (2018). Data analytics: The future of audit. IBR.

De Bonhome, O., Jans, M., Marissen, M., Simpelaere, F., Trumpener, J. & Verachtert, S. (2020). Het gebruik van data-analyse bij de controle. IBR.

De Weerdt, J., De Backer, M., Vanthienen, J. & Baesens, B. (2012). A multidimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs. *Information Systems*, *37*(7), 654–676.

Depaire, B., Swinnen, J., Jans, M. & Vanhoof, K. (2013). A Process Deviation Analysis Framework, In *International Conference on Business Process Management*, Berlin, Heidelberg, Springer.

Dikici, A., Turetken, O. & Demirors, O. (2018). Factors influencing the understandability of process models: A systematic literature review. *Information and Software Technology*, *93*, 112–129.

Flores, G. & Riquenes, R. (2020). Audit 2.0, a perspective for its execution in the business environment using process mining techniques. *Vivat Academia*, *23*(150), 47–57.

GDPR one year on: Its impact on auditors and accountants? (2019). Retrieved March 27, 2021, from https://www.accountancyeurope.eu/audit/gdpr-one-year-on-its-impact-on-auditors-and-accountants/

GDPR: Implications for auditors. (2018). Retrieved March 27, 2021, from https://www.accountancyeurope.eu/publications/gdpr-implications-for-auditors/

Genga, L., Alizadeh, M., Potena, D., Diamantini, C. & Zannone, N. (2018). Discovering anomalous frequent patterns from partially ordered event logs. *Journal of Intelligent Information Systems*, *51*(2), 257–300.

Gillin, R. & Spoor, E. (2018). GDPR issues for auditors. Retrieved March 27, 2021, from https://www.icaew.com/technical/audit-and-assurance/faculty/webcasts/gdpr-issues-for-auditors

Goedertier, S., Martens, D., Vanthienen, J. & Baesens, B. (2009). Robust Process Discovery with Artificial Negative Events. *Journal of Machine Learning Research*, *10*(44), 1305–1340.

Günther, C. W. & van der Aalst, W. M. P. (2007). Fuzzy Mining – Adaptive Process Simplification Based on Multi-perspective Metrics, In *International conference on business process management*, Berlin, Heidelberg.

Jans, M., Alles, M. & Vasarhelyi, M. (2010). *Process Mining of Event Logs in Auditing: Opportunities and Challenges* (SSRN No. 1578912). Social Science Research Network. Rochester, NY.

Jans, M., Alles, M. & Vasarhelyi, M. (2013). The case for process mining in auditing: Sources of value added and areas of application. *International Journal of Accounting Information Systems*, *14*(1), 1–20.

Jans, M., Alles, M. & Vasarhelyi, M. (2014). A Field Study on the Use of Process Mining of Event Logs as an Analytical Procedure in Auditing. *The Accounting Review*, *89*(5), 1751–1773.

Jans, M. & Hosseinpour, M. (2019). How active learning and process mining can act as Continuous Auditing catalyst. *International Journal of Accounting Information Systems*, *32*, 44–58.

Janssenswillen, G., Donders, N., Jouck, T. & Depaire, B. (2017). A comparative study of existing quality measures for process discovery. *Information Systems*, *71*, 1–15.

Leemans, M. & van der Aalst, W. M. P. (2015). Discovery of Frequent Episodes in Event Logs, In *International symposium on data-driven process discovery and analysis*, Cham, Springer.

Leemans, S. J. J., Fahland, D. & van der Aalst, W. M. P. (2013). Discovering Block-Structured Process Models from Event Logs - A Constructive Approach, In *International conference on applications and theory of Petri nets and concurrency*, Berlin, Heidelberg, Springer.

Leemans, S. J. J., Fahland, D. & van der Aalst, W. M. P. (2014). Discovering Block-Structured Process Models from Event Logs Containing Infrequent Behaviour, In *International conference on business process management*, Cham, Springer.

Li, W., Zhu, H., Liu, W., Chen, D., Jiang, J. & Jin, Q. (2018). An Anti-Noise Process Mining Algorithm Based on Minimum Spanning Tree Clustering. *IEEE Access*, *6*, 48756–48764.

Măruşter, L., Weijters, A. T., van der Aalst, W. M. P. & Van Den Bosch, A. (2006). A Rule-Based Approach for Process Discovery: Dealing with Noise and Imbalance in Process Logs. *Data Mining and Knowledge Discovery*, *13*(1), 67–87.

Mitsyuk, A. A. & Shugurov, I. S. (2016). On process model synthesis based on event logs with noise. *Automatic Control and Computer Sciences*, *50*(7), 460–470.

Nguyen, H. T. C., Lee, S., Kim, J., Ko, J. & Comuzzi, M. (2019). Autoencoders for improving quality of process event logs. *Expert Systems with Applications*, *131*, 132–147.

Polančič, G. & Cegnar, B. (2017). Complexity metrics for process models – A systematic literature review. *Computer Standards & Interfaces*, *51*, 104–117.

Putra, L. D. (2009). 7 Major Phases of Audit [of Financial Statements]. Retrieved October 4, 2020, from http://accounting-financial-tax.com/2009/09/7-major-phases-of-audit-of-financial-statements/

Suriadi, S., Andrews, R., ter Hofstede, A. H. M. & Wynn, M. T. (2017). Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. *Information Systems*, *64*, 132–150.

Syring, A. F., Tax, N. & van der Aalst, W. M. P. (2019). Evaluating Conformance Measures in Process Mining Using Conformance Propositions, In *Transactions on Petri Nets and Other Models of Concurrency XIV*. Berlin, Heidelberg, Springer.

Tax, N., Lu, X., Sidorova, N., Fahland, D. & van der Aalst, W. M. P. (2018). The Imprecisions of Precision Measures in Process Mining. *Information Processing Letters*, *135*, 1–8.

Tax, N., Sidorova, N., Haakma, R. & van der Aalst, W. M. P. (2016). Mining local process models. *Journal of Innovation in Digital Ecosystems*, *3*(2), 183–196.

Theis, J. & Darabi, H. (2020). Adversarial System Variant Approximation to Quantify Process Model Generalization. *IEEE Access*, *8*, 194410–194427.

van den Broucke, S. K. L. M., De Weerdt, J., Vanthienen, J. & Baesens, B. (2013). A comprehensive benchmarking framework (CoBeFra) for conformance analysis between procedural process models and event logs in ProM, In *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*.

van den Broucke, S. K. L. M., De Weerdt, J., Vanthienen, J. & Baesens, B. (2014). Determining Process Model Precision and Generalization with Weighted Artificial Negative Events. *IEEE Transactions on Knowledge and Data Engineering*, *26*(8), 1877–1889.

van der Aalst, W. (2016a). 3.2: On The Representational Bias of Process Mining - Different Types of Process Models. Retrieved December 28, 2020, from https://www.coursera.org/lecture/process-mining/3-2-on-the-representational-bias-of-process-mining-HPJUZ

van der Aalst, W. (2016b). *Process Mining: Data Science in Action (Second Edition)*. Springer-Verlag Berlin Heidelberg.

van der Aalst, W., Adriansyah, A. & van Dongen, B. (2012). Replaying history on process models for conformance checking and performance analysis. *WIREs Data Mining and Knowledge Discovery*, *2*(2), 182–192.

van Zelst, S. J., van der Aalst, W. M. P. & Verbeek, H. M. W. (2018). Discovering workflow nets using integer linear programming. *Computing*, *100*(5), 529–556.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R. & Cleven, A. (2015). Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research. *Communications of the Association for Information Systems*, *37*(1), 9.

Werner, M. (2017). Financial process mining - Accounting data structure dependent control flow inference. *International Journal of Accounting Information Systems*, *25*, 57–80.

Wiese, G. (2019). Thinking outside the SOX: 3 simple steps to ensure "IPE" SOX Compliance. Retrieved October 2, 2020, from https://www.pwc.co.uk/risk-assurance/assets/thinking-outside-sox-ensure-ipe-compliance.pdf

Wynn, M. T. & Sadiq, S. (2019). Responsible Process Mining - A Data Quality Perspective, In *17th International Conference on Business Process Management*, Vienna, Austria.

Zhang, J. (2017). A logic-based representation and tree-based visualization method for building regulatory requirements. *Visualization in Engineering*, *5*(1), 2.

9. Appendices

9.1. Appendix 1: Metrics for process discovery requirements in audit practice

Requirement	Evaluation based on algorithm's publication	Recommended metrics	Motivation
Fitness		→ Alignment-based fitness → Negative event recall → Eigenvalue fitness	Recommended by comparative studies such as Janssenswillen et al. (2017) and Syring et al. (2019).
Approach to noise	V		Solely on-the-spot measures of the amount of noise are available.
Precision		→ One-align precision → Alignment-based precision → Anti-alignment based precision → Eigenvalue precision → Soundness	Recommended by comparative studies such as Janssenswillen et al. (2017), Syring et al. (2019) and Tax et al. (2018).
Generalisation	V		There is little consensus between existing metrics.
Ability to deal with log quality problems	V		Check for measures taken to ensure log quality.
Simplicity	V	Model size, unstructuredness, control flow complexity, entropy... (overview in Polančič and Cegnar (2017))	Create an all-inclusive, befitting set of metrics or check for measures taken to lower structural complexity.
Practicality		→ Run time → Memory use	
Approach to materiality	-	-	Currently not implemented.

9.2. Appendix 2: Plug-ins applied in the demonstration

The table below lists the various ProM[2] plug-ins, CoBeFra[3] metrics, and Apromore[4] functionalities used in the demonstration. On every occasion, default parameter values were used, as parameter optimisation was considered to be outside the scope of this demonstration. When required for the metrics' calculation, the miner outputs were converted to Petri nets.

Action	**Tool**	**Plug-in or functionality**	**Package**
Mine process model: Inductive Miner	ProM	Mine Petri net with inductive miner	InductiveMiner
Mine process model: Evolutionary Tree Miner	ProM	Mine a Process Tree with ETMd	Evolutionary-TreeMiner
Convert process tree to Petri net (for ETM)	ProM	Convert Process Tree to Petri Net	/
Mine process model: Split Miner	Apromore	Discover model (BPMN)	/
Convert BPMN model to Petri net (for Split Miner)	ProM	Convert BPMN diagram to Petri net (option: 'translate for: conformance checking')	/
Mine process model: Integer Linear Programming Miner	ProM	ILP-Based Process Discovery	HybridILPMiner
Calculate fitness metric: Alignment-based fitness	ProM	Replay a Log on Petri Net for Conformance Analysis	PNetReplayer
Calculate fitness metric: Negative event recall	CoBeFra	Negative event recall	/
Calculate precision metric: One-align precision	ProM	Check precision based on Align-ETConformance	ETConformance
Calculate precision metric: Alignment-based precision	CoBeFra	Alignment Based Precision	/

[2]http://www.promtools.org/ (version 6.10)

[3]http://processmining.be/cobefra/ (version of 2018-04-03), (van den Broucke et al., 2013)

[4]https://apromore.org/