

## Process Mining in Auditing: From Current Limitations to Future Challenges

Mieke Jans

Hasselt University  
Faculty of Business Economics  
3590 Diepenbeek  
Belgium  
mieke.jans@uhasselt.be

In the first book on process mining, Wil van de Aalst densely defines the goal of process mining "to use event data to extract process-related information", like automatically discovering a process model by observing events that are recorded by some information system. [1] This definition is broad, since it addresses the mining of all processes that are supported by an information system, revealing the wide range of possible applications of process mining. With the growing of the digital universe, the recording of events reaches new heights all the time. Events can relate to scanning goods at the supermarket, filling out an online test for the best offer on mobile phones, buying an e-ticket for your next conference, or using the newest app on your smart phone. Events are also recorded in a business context, ranging from a stand-alone sale to all related events in a Customer Relationship Management system.

Given the omnipresence of recorded events in the context of a process and hence the large amount of possibilities to apply process mining, a well-defined focus on an application field is essential. Auditing is such a field. The auditor functions as an independent examiner of financial statements to give reasonable assurance on the accuracy of these statements. That way, the auditor provides 'trust' to shareholders and other third parties related to the audited organization. This trust is a crucial element of the economic system. Indeed, if one could not rely on financial statements because of a lack of trust, no one would be willing to provide capital to organizations, posing a true threat to the continuation of these organizations.

The profession of auditing is overseen by international and national non-profit organizations. The two largest oversight bodies are the Public Company Accounting Oversight Board (PCAOB) in the US and the International Auditing and Assurance Standards Board (IAASB). Auditing standards like the ones issued by the PCAOB and the IAASB place an important emphasis of

auditing on understanding the processes that precede financial reporting. In order to understand the likely sources of misstatement, the auditor has to understand the flow of transactions and to identify the controls that management has implemented to address potential misstatements or to prevent unauthorized acquisition, use, or disposition of a company's assets (AS No. 5, paragraph 34). To achieve these objectives, the standards present the use of walkthroughs as the most effective means. Walkthroughs are performed by following a transaction from origination through a process, including information systems, until it is reflected in the company's financial records. This approach, currently used in the auditing profession, can be drastically ameliorated by employing the techniques of process mining in order to achieve abovementioned objectives. The technique of the walkthroughs can, by applying process mining, (1) be automated, and (2) extended to the full population instead of a sample, resulting in a transparent overview of the process. However, in order to fully replace the technique of manual walkthroughs by process mining, several current limitations need to be addressed. These limitations are both practical and theoretical. These limitations are summarized in the following paragraphs, starting from the most theoretical to the most practical limitations.

An important aspect of an audit, is that it should provide reasonable assurance about whether the financial statements are free of material misstatement. A fact is defined as 'material' if there is "a substantial likelihood that the fact would have been viewed by the reasonable investor as having significantly altered the 'total mix' of information made available." (AS No. 11, paragraph 2) This dimension of materiality requires a delicate assessment of the auditor, taking into account all information and all possible misstatements in order to state whether a level of materiality is reached or not. Hence this is not easy to convert into exact rules. Therefore, a conciliation between the definition of materiality and a process mining approach should be found.

In the search for a fit between materiality and process executions, it is important to keep in mind the different dimension a process represents when compared to an account. A process execution, or pattern, is a path that can be followed and that ends in crediting or debiting an account. A process execution at itself cannot give insight into materiality. To this end, extra information needs to be added to the pattern, like how many transactions followed this path, which value was created by following this path, how many persons were involved, does the path cover more than one financial booking period? As such, questions that need to be answered in the pursuit of a translation of processes to materiality, are amongst other:

- When is a certain deviating process execution warranting further examination to exclude a material misstatement?

- How can a process deviation be quantified in terms of risk?
- Is there a certain threshold of cases that follow a certain process execution to consider it as material, or a threshold on the affected amount of money?

To answer these questions, maybe new algorithms with new, better suited perspectives need to be developed. The materiality issue is an important issue in auditing and given its delicate assessment, this probably will not be answered very easily. Probably this issue will in a process mining approach, just like in the current approach, require expertise of the auditor and cannot be fully replaced by algorithms. This however does not refrain us from searching for indicators to help the auditor in his task.

Aside from the materiality issue, the process selection requires our attention. In order to mine the processes that result in financial reporting, one needs to identify the relevant processes. On the one hand, there are the traditional business cycles, like procurement and sales, that result in financial reporting output. On the other hand, each company might have its company specific processes affecting the financial statements. For process mining to be fully integrated as part of auditing, it is important to identify the relevant processes, taking the organization's structure as a starting point. This results in a completely customized approach for each audit. One approach to deal with this element is to make this assessment a part of the risk assessment phase of an audit.

To supersede the entire range of the manual walkthroughs, it is important that the process under examination is completely embedded in an information system. The start transaction, all following transactions, and the final financial reporting transaction need to be captured by the information system. If not, the automated process mining will not be able to mine the whole process, but only the part that is supported by an information system. This restriction demands a certain level of maturity of the organization, before process mining can be applied in an auditing context. If only part of a process can be mined, no assurance can be provided on the process and its reporting outcome. The consequence of this limitation is that a full integration for all audits cannot be reached. However, since the digitalization of the world continues, more and more organizations and processes will be suited for process mining. If this trend continues, as assumed, the application possibilities of process mining will grow each day. By the time the process mining field is fully adopted to audit requirements, most audits will take place in an information system environment and can be complemented and improved by process mining techniques. However, the official auditing standards will always have to make a distinction between processes or organizations that are suited for process mining (and hence are subjected to more thorough examinations) and those that are not. If not framed correctly, this distinction may turn into a breeding ground for discussions and/or abuse. This issue has to be thought

over seriously when inserting process mining into standards.

In order to apply process mining techniques, some conditions concerning the data structure must be met. These conditions form a current limitation on experiencing the full advantage of today's process mining capabilities. The main issue is that advanced ERP systems are based on a relational database structure. Process mining on the other hand starts from an event log where all activities take place on the same level of process instances. An example makes this clear. The information concerning an invoice might be captured in two tables of an ERP system: a table containing header information and a table containing item line details. Events that relate to this invoice may affect the header table or the item detail table. For example the approval of the invoice will occur at header level, while the booking of this invoice will take place on line item level. These two levels make it difficult to assign process activities to one process instance. If an invoice -as a whole- is selected as process instance to follow throughout the process and its subsequent activities, it is not possible to accurately link the activities on item level -like booking the separate item lines- to the process instance. On the other hand, if the item level detail were chosen as process instance, the activities on header level -like approval- are not correctly matched. One activity on header level would be presented multiple times in all related item level process instances. It still is possible to mine the process with these limitations, but it creates extra manual examination of the output, exactly what process mining aims to reduce. This extra examination can be discarded by finding a solution for the problem of non-compliant data structures in the field of process mining on the one hand and in the most used ERP systems on the other hand. This solution holds either an evolution of process mining into relational process mining, or a structural change in the ERP systems regarding the link between activities and levels. Both evolutions are not straightforward or expected in the very near future, making this issue of data structure one of the most important threats of the adoption of process mining in auditing.

In case process mining in auditing is naturalized, evidentially speaking qualitative tools for both event log building and process mining analysis are needed. There are already some providers of commercial process mining tools and one provider of an event log building tool. However, more work needs to be done, not the least in adapting these general tools to specific audit requirements. The IT capabilities of the end user, the auditor, has to be taken into account; the typical audit related questions and interests; the type of desired output; and the related responsibilities. Also securing the extracted data out of the information system is an important issue and needs to be dealt with. One option is to secure the data like in some Generalized Audit Software where no alterations are possible once the data is imported into the tool.

Another point of interest are the identified outliers. In case an auditor identifies outliers -in accounts, transactions or processes-, the auditor needs to clear these outliers of containing a risk of material misstatements. The positive identification of outliers in process executions are true positives, as opposed to some data mining algorithms that might also output false positives. However, the current, existing process discovery algorithms in the domain of process mining are not yet adapted to auditing purposes, since the discovered patterns do not report in a comprehensive way on loops. This gives the impression that hundreds of different process executions exist, while in reality it might be a combination of loops. New algorithms that output 'head'-patterns along with the identification of 'sub'-patterns that are possibly repeated within one pattern need to be created.

As a last item, I would like to point out that process mining increases the insights in the examined process and there is a possibility that an auditor will experience this as an increasing work factor. It is important to counter this perception. These better insights that are gained are a leap forward in the assurance that the auditor, both internal and external, can provide on an organizations financial statements. Today's audit will not only be automated, but will be drastically reengineered. The efforts to adapt the current domain of process mining to audit requirements and possibly also vice versa are truly necessary for an adoption success. But this would be in a later stadium, after the adaptation of the process mining field to the auditing profession. Of course, since process mining is only suitable to (exa)mine processes that are fully embedded in information systems, the application of process mining will always be restricted to those organizations with a sufficient maturity, making the 'process mining standards' deviating from the regular audit standards, which are applicable to all firms.

As an overall conclusion we can say auditing is an interesting field that can benefit from what process mining has to offer. The needs of auditing are complement to the aims of process mining: clarifying processes with the intent of providing assurance. This complement aspect should encourage both researchers and practitioners to seek for solutions on the current limitations. The limitations require further fundamental research on algorithms and applied research on practical process mining implementations to learn from. In the end, the standard setting bodies need these inputs to present a process mining approach as part of the auditing examination.

## References

1. van der Aalst, Wil: Process Mining: Discovery, Conformance, and Enhancement of Business Processes, Springer (2011)