Data analytics in external auditing

DISSERTATION
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Management

submitted by

Anita Stefanie Gierbl

from

Austria

Approved on the application of

Prof. Dr. Peter Leibfried

and

Prof. Dr. Peter Rohner

Dissertation no. 5044

Difo-Druck GmbH, Untersiemau 2021

The University of St.Gallen, School of Management, Economics, Law, Social Sciences, International Affairs and Computer Science hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St.Gallen, October 23, 2020

The President:

Prof. Dr. Bernhard Ehrenzeller

Acknowledgments

I would like to express my gratitude to everyone who has supported and accompanied me on my dissertation journey.

First and foremost, I am grateful to my PhD supervisor Prof. Dr. Peter Leibfried for giving me the opportunity to write this dissertation, especially on such a popular topic, and for his ongoing support during this process. Through him, I had the great honor to work for Swiss GAAP FER as a technical assistant, where I had the pleasure to meet and work with brilliant people. As such, I would like to extend my gratitude to everyone at Swiss GAAP FER.

I would further like to thank Prof. Dr. Peter Rohner for taking the role of my co-supervisor for this dissertation. Thanks to his valuable advice, especially during the discussion of my preliminary study, this dissertation received valuable nudges in the right direction.

Moreover, I would like to express my sincere gratitude to everyone I interviewed for this dissertation. It was a great pleasure to talk with everybody about this exciting research topic and to receive many brilliant insights and inputs.

A big thanks goes out to all my friends for their friendship and support, and especially to those who reviewed this dissertation, for their highly appreciated and valuable feedback.

I am also very grateful for the continued support, motivation, and patience of my beloved partner. Finally, my deepest gratitude goes to my parents Maria and Josef Gierbl.

Zurich, February 2020 Anita Gierbl

Summary

This dissertation examines the use of data analytics in external auditing. Research findings, current practice, and future applications are investigated to evaluate the potential of data analytics for auditing.

The findings indicate that "conventional data analytics" methods (rule-based queries, visualization, and descriptive statistics) are widely used in the auditing industry today, while "advanced data analytics" approaches (machine learning, process mining and natural language processing) are in its infancy stage. The primary exceptions are process mining and natural language processing which are enjoying some early successes.

Examples of future use cases in the field of machine learning are journal entry testing (clustering, adversarial autoencoder neural networks, association rules), predictions for analytical procedures, and audit workforce planning.

Overall, external auditing is a promising field for the application of data analytics. Successful adoption will have implications on the audit market (e.g., adoption of data driven audits in phases), the audit profession (e.g., extension of skillset) and the regulator (e.g., ensuring adequate application).

Zusammenfassung

Die vorliegende Dissertation befasst sich mit Einsatzmöglichkeiten von Datenanalyseverfahren (Data Analytics) in der Wirtschaftsprüfung. Um das Potential von Data Analytics für die Prüfung zu beurteilen, werden drei Bereiche untersucht: Forschungserkenntnisse, derzeitige Praxis und zukünftige Anwendungsmöglichkeiten.

Basierend auf den Ergebnissen werden "konventionelle Datenanalyse-methoden" (regelbasierte Abfragen, Visualisierung und deskriptive Statistik) bereits weitgehend in den untersuchten Unternehmen eingesetzt. Fortgeschrittene Datenanalyseverfahren (Machine Learning, Process Mining und Natural Language Processing) finden in der derzeitigen Prüfungspraxis noch kaum Anwendung. Process Mining und Natural Language Processing hingegen können bereits erste Erfolge vorweisen.

Beispiele für zukünftige Anwendungen im Bereich Machine Learning sind: Journal Entry Testing (Clustering, Adversarial Autoencoder Neural Network, Assoziationsanalyse), Erwartungswerte für analytische Prüfungsverfahren und Personaleinsatzplanung.

In der Dissertation wird aufgezeigt, dass ein breites Anwendungsgebiet für Datenanalyseverfahren in der Wirtschaftsprüfung besteht. Dies führt zu vielfältigen Auswirkungen für den Prüfungsmarkt (z.B. Adaptionsphasen des Data Driven Audits), Berufsstand (z.B. Veränderung des Kompetenz-spektrums) und Regulator (z.B. Sicherstellung einer angemessenen Anwendung).

Content overview

Table of contents

List of figures

List of tables

List of abbreviations

ACFE	Association of Certified Fraud Examiners
ACL	Audit command language
AI	Artificial intelligence
ARCH	Autoregressive conditional heteroscedasticity
ARIMA	Autoregressive integrated moving average
AUC	Area under the curve
CAAT	Computer assisted audit techniques
CFO	Chief Financial Officer
CHF	Swiss francs (currency)
COSO	Committee of Sponsoring Organizations of the Treadway Commission
CPA	Certified public accountant
CRISP-DM	Cross Industry Standard Process for Data Mining
ERP-system	Enterprise resource planning system
FAOA	Swiss Federal Audit Oversight Authority
GARCH	Generalized autoregressive conditional heteroscedasticity
IAASB	International Auditing and Assurance Standards Board
IDEA	Interactive Data Extraction and Analysis
IFRS	International Financial Reporting Standards
IAS	International Accounting Standard
ISA	International Standard on Auditing
ISQC	International Standard on Quality Control

ISQM	International Standard on Quality Management
IT	Information technology
k-nn	k-nearest neighbor
KPI	Key performance indicators
LOF	Local outlier factors
MD&A	Management discussion and analysis
ML	Machine learning
n.d.	no date
NLP	Natural language processing
OCR	Optical character recognition
ProM	Open source process mining software for process mining
PS	Swiss Auditing Standards (Prüfungsstandard)
RVI	Relative variable importance
SQL	Structured query language
SVM	Support vector machines
US-GAAS	United States Generally Accepted Auditing Standards
VHB	Verband der Hochschullehrer für Betriebswirtschaft

1 Introduction

1.1 Research relevance and motivation

We are at the start of the 4th industrial revolution, in which technology and digitalization will fundamentally transform the way we live, work, and relate to one another (Schwab, 2017). Billions of people are connected to the internet via a multitude of devices, opening new opportunities for constant access to knowledge and communication (Schwab, 2017). Processing power, storage capabilities, and the speed at which we can access information are reaching new all-time highs every year. We also have the software which allows us to analyze massive amounts of data. All of this gives rise to the hype and media coverage about big data and artificial intelligence.

Artificial intelligence, in particular, is experiencing a second renaissance. Although the technology first emerged in the 1940s (Haenlein & Kaplan, 2019), the required processing power, storage capacity, and broad availability of data only recently caught up, reaching a level that allows the technology and its underlying disciplines, such as machine learning, to reach their potential.

The emergence of technologies creates new business models and jobs. Meanwhile, established business models and professions become obsolete, or are at least reshaped, in an active process of creative destruction (Schumpeter, 1942). This has far-reaching implications for many professions that long appeared to be safe from digitalization and automation. An extreme example comes from the medical field. Although humans spend several years in medical school, recent results show that artificial intelligence has a better classification success rate when it comes to diagnosing skin cancer through malignant melanomas (Donnelly, 2018; Lindner, 2019). If, in a sector with such a welleducated workforce, artificial intelligence can be used and performs better than highly-trained humans, is any profession safe from automation?

Frey and Osborne (2017) conducted a study focused on answering this exact question. They analyzed 702 occupations in detail and estimated the probability of automation for every single one. According to their results, recreational therapists' jobs are least likely to be automated, while telemarketers have the highest probability of being replaced by machines. Also high on the automation scale is auditing, with a probability of 94%. The authors' estimated timeframe for this change is within the next ten to twenty years. Many senior technology and communications executives expect that 30% of all audits will be performed by artificial intelligence within the next decade (World Economic Forum, 2015). The fact that technologies such as artificial intelligence are expected to play an essential role in the future of auditing is also shown by the survey of Burrus (2018), in which more than 1'000 US auditors, finance professionals, and accounting professionals participated.

Issa, Ting, and Vasarhelyi (2016) compare the audit process to a production line, the product of which is an opinion about the audited financial statements based on professional judgment. At the moment, this audit process is a highly labor-intensive service produced in a hierarchical structure. Issa et al. (2016, p. 12) describe this structure as "lower-level employees performing repetitive low-level tasks (ticking, extending) in verifying manual documents, and their hierarchical superiors examining (reviewing) these tasks and drawing conclusions, which are typically judgmental". Such a structured and labor-intensive service with so much grunt work seems fertile ground for automation.

The audit companies known as the "Big 4" (Deloitte, EY, KPMG, and PwC) are addressing this topic individually or have formed collaborative partnerships to incorporate more technology into the audit process. Deloitte developed an advanced data analytics tool named Argus to extract relevant information from any document (Deloitte, 2016). EY created an audit data analytics toolbox called Helix, which includes a capability for analyzing sales invoicing activities throughout the year (Ernst & Young, 2019). KPMG developed a tool named Clara, which incorporates data analytics into the audit process (KPMG, 2017). PwC is teaming up with H2O.ai[1] to detect anomalies in general ledgers with artificial intelligence (PwC, 2019). All these innovative tools have one thing in common: they apply data analytics to auditing using various methods, including artificial intelligence.

Incorporating such tools promises to reduce the resources required to conduct an audit (efficiency) while simultaneously gaining a higher level of audit assurance (effectivity) and consequently lowering the costs of an audit (Eberle, 2017). No scientific evidence was found that any of the named benefits have been realized to date. Still, the results of a survey conducted among Swiss companies show that 34% expect audit fees to decrease in the future due to the digitalization of the audit industry (Fontana, Birrer, Hummel, & Füchslin, 2018).

Nevertheless, the history of auditing and technology is rife with failed implementations and unfulfilled promises. Alles and Gray (2016) report that around 25 years ago, overly optimistic articles were published on various technologies considered useful for the auditing profession. In the 80s and 90s, an extensive research community was focused on how expert systems and neural networks, an approach to artificial intelligence, could be used for external audit tasks (Omoteso, 2012). The anticipated disruption, however, failed to materialize in practice. While some of the results of that period's research have since been embedded in audit tools, they failed to yield stand-alone applications for financial statement audits (Dowling & Leech, 2007).

[1] H20.ai is a US-based artificial intelligence company.

Fast forward to today, however, the situation appears more promising. Processing power and storage capacity have caught up, and massive amounts of data are now available. In recent decades companies have invested extensively in IT infrastructure throughout their operational processes (Provost & Fawcett, 2013). This transformation has resulted in more audit-relevant data being stored electronically and digitally available. Thus, audit companies are expecting that data analytics, in its many forms and incorporating various technologies, will transform auditing (e.g., Deloitte, 2016; KPMG, 2017). Huerta and Jensen (2017) believe that data analytics has the potential to automate auditing procedures.

Will this time be different? And if it is, what are the implications? The motivation of this dissertation is to address both questions by analyzing the use of data analytics in external auditing and what it means for the audit market, audit profession, and regulator.

1.2 Research questions

A significant change in auditing is currently anticipated due to technological developments. Driven by technological innovations of the Big 4 companies, data analytics has emerged as a central topic for the auditor. Research also predicts that data analytics will have a significant influence on auditing (Appelbaum, Kogan, & Vasarhelyi, 2017; Appelbaum, Kogan, & Vasarhelyi, 2018; Earley, 2015; Schneider, Jun, Janvrin, Ajayi, & Raschke, 2015; Titera, 2013; Wang & Cuthbertson, 2014).

However, research gaps exist in regards to the use of data analytics in external auditing (Appelbaum et al., 2018; Earley, 2015; Koh & Low, 2004; Wang & Cuthbertson, 2014). Particularly, empirical research on this topic is in its infancy (Salijeni, Samsonova-Taddei, & Turley, 2019).

The following research questions were developed based on identified research gaps. As each research question dictates the appropriate research method (Eriksson & Kovalainen, 2016), each applied research design is briefly presented.

While Appelbaum et al. (2018) carried out a comprehensive literature review regarding data analytics in auditing, they did not focus on the practical applications of each data analytics method with real-world data. Hence, the first research question of this dissertation is an extension and specification of their paper and asks:

Research question 1: Which applications of data analytics methods have been identified for external auditing in research?

A structured literature review answers this question. Since processing power and storage capacity have expanded exponentially over time, only publications from 2000 onwards are considered. 50 journals and three electronic databases were searched to identify the relevant literature.

Earley (2015) and Appelbaum et al. (2018) conclude that a research gap exists on the data analytics tools currently used by practitioners. Consequently, the second research question is:

Research question 2: Which data analytics methods have been implemented in the external audit process?

Interviews are conducted with eight Swiss audit companies to answer this research question. The audit process is used as a framework to put the identified applications in context.

Having identified data analytics use cases within the literature and the current applications in auditing, further potential applications of data analytics for external auditing are considered in collaboration with technical experts. This yields the next research question:

Research question 3: What are potential future applications for data analytics methods in external auditing?

To answer this research question, 13 interviews with data analytics experts are conducted. Based on the interviews, examples of future applications of data analytics methods for external auditing are identified.

The final research question is devoted to the call for research by Wang and Cuthbertson (2014) regarding the implications of using data analytics in auditing, with a focus on the audit market, audit profession, and regulator:

Research question 4: What are the implications of using data analytics in external auditing for the audit market, audit profession, and regulator?

The implications build on the findings of the preceding research questions and the corresponding collected data. The implications identified by the author were further validated during six interviews.

1.3 Structure of the dissertation

Figure 1 shows the structure of this dissertation.

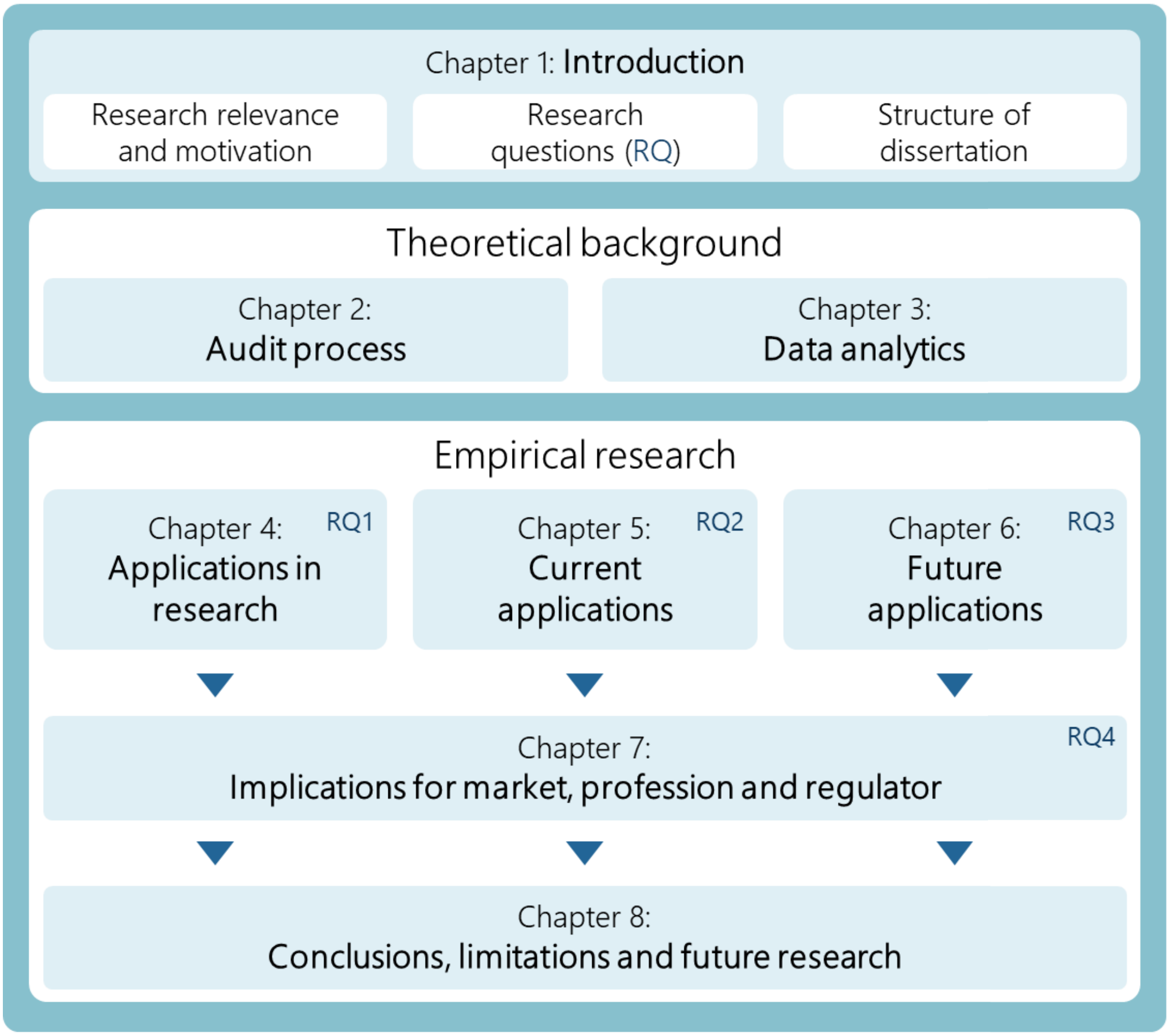

Figure 1 Structure of the dissertation

Chapter 1 begins with an outline of the research relevance and motivation, leading to the identified research gaps and research questions of this dissertation.

Chapters 2 and 3 provide the necessary theoretical background.

Chapter 2 briefly depicts the whole audit process and explains which tasks auditors perform during a full statutory audit based on the International

Standard on Auditing (ISA). The complete audit process is visualized in order to help identify where data analytics can be applied within auditing.

Chapter 3 explains the process of performing data analytics. Within this chapter, the focus is on data analytics methods, as these are essential to understanding the applications presented in the later chapters.

Subsequently, each research question is addressed in a single chapter. The outline of each empirical chapter is the same. It starts with an introduction of the research design, followed by a summary of the sample selection, data collection, data analysis phase and presentation of the results.

Chapter 4 (research question 1) presents the identified uses cases of data analytics in the literature and discusses the findings.

Chapter 5 (research question 2) focuses on data analytics applications currently used in audit practice. The first outline of the results presents currently applied data analytics methods and tools, which are then mapped onto the audit process. After that, motivations for applying data analytics, as well as current challenges, are addressed. Lastly, the role of the standard setter regarding the application of data analytics is discussed.

Chapter 6 (research question 3) presents and reflects on five potential future applications, based on interviews conducted with experts.

Chapter 4, 5 and 6 end with a summary of the findings and the corresponding impact on the audit process.

Chapter 7 (research question 4) focuses on the implications of using data analytics in external auditing on the audit market, audit profession, and regulator.

Chapter 8 concludes the dissertation by summarizing the findings of all research questions. It furthermore addresses the limitations of this dissertation and provides an outlook for future research.

2 Audit process

*«The **purpose** of an audit is to **enhance the degree of confidence** of intended users in the financial statements. This is achieved by the expression of an **opinion** by the auditor on whether the **financial statements are prepared, in all material respects, in accordance with an applicable financial reporting framework**» (ISA 200.3; emphasis added)*

At the end of an audit, the auditor expresses an opinion about the overall accuracy of financial statements according to the applied accounting standard, as required by ISA 200.3. In order to come to a conclusion about the financial statements, the auditor collects evidence during the audit. According to ISA 500.5c "audit evidence includes both information contained in the accounting records underlying the financial statements and other information".

The auditor is required to conduct the audit within the parameters of regulatory frameworks (e.g., Swiss Code of Obligations) and professional standards, which makes the profession complex (Appelbaum et al., 2017). Various auditing standards, such as the International Standards on Auditing (ISA), the United States Generally Accepted Auditing Standards (US-GAAS), or the Swiss Auditing Standards (PS[2]), define the audit process. This dissertation focuses on the ISA standard, it being the most relevant standard from a European perspective. The fact that the Swiss Auditing Standard has largely implemented the ISA standard[3] (Schneider & Bradtke, 2017) allows this literature also to be used in the depiction of the audit process.

Once the audit engagement is accepted, the audit process starts with audit planning and ends with the handing over of the final audit report. The whole process can be divided into three major stages: audit planning, audit

[2] Prüfungsstandard

[3] The Swiss Auditing Standard also contains special auditing standards such as limited statutory examination, which are not discussed in the following account of the audit process.

performance, and audit completion (Arens, Elder, Beasley, & Hogan, 2017; EXPERTsuisse, 2015). In parallel to the three stages, two supporting procedures – documentation and quality control – are carried out continuously during the audit. Issa et al. (2016, p. 11) compare the flow of the audit process to an assembly line, where "the output of one phase becomes the input of the subsequent one", and the final product is the opinion in the audit report. In practice, however, the audit process is often more of an iterative process than a strictly linear one, where the auditor may need to adjust some input parameters (e.g., risk assessment) from the previous stage (ISA 300.10).

The term "auditor" within this dissertation refers to the engagement partner (ISA 200.13d). A high degree of individual, professional, and expert judgment is required of the auditor during an audit (Salijeni et al., 2019). In this regard, the terms "professional judgment" and "professional skepticism" are frequently mentioned. Professional judgment is defined in ISA 200.13k as "the application of relevant training, knowledge and experience, within the context provided by auditing, accounting and ethical standards, in making informed decisions about the courses of action that are appropriate in the circumstances of the audit engagement". Professional skepticism is described as "an attitude that includes a questioning mind, being alert to conditions which may indicate possible misstatement due to error or fraud, and a critical assessment of audit evidence" (ISA 200.13l).

Figure 2 shows the audit process of a full statutory audit, which is described in the following sections.

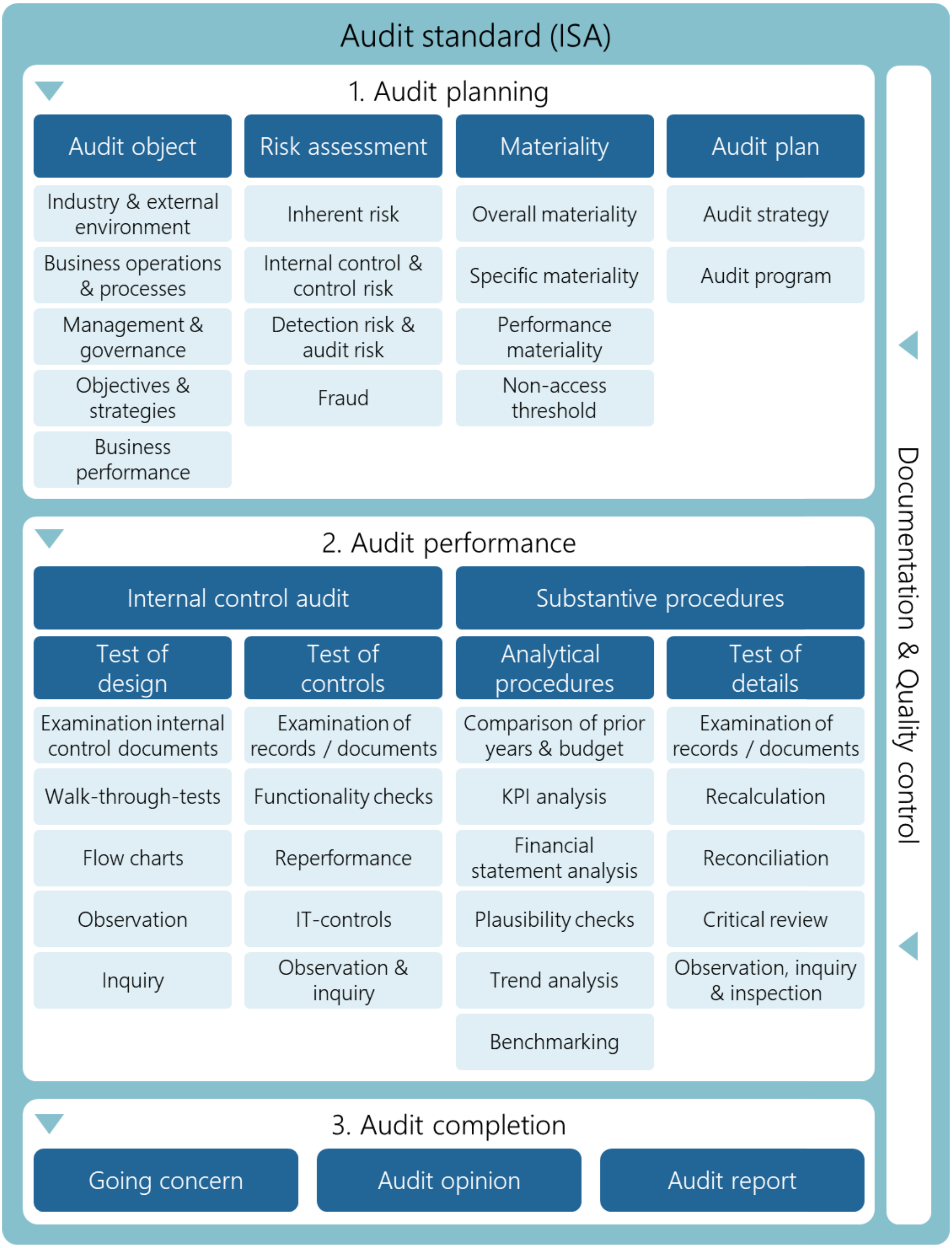

Figure 2 Audit process based on Arens et al. (2017); EXPERTsuisse (2015); IAASB (2019a); Marten, Quick, and Ruhnke (2015)

2.1 Audit planning

The first stage of the audit process is audit planning. "The nature and extent of the planning" (ISA 300.A1) depends on the size of the audit client (audit object), the company's complexity in terms of activities, industry, and business environment, as well as the auditor's experience with the company (Kartscher, Rossi, & Suter, 2013). An aircraft manufacturer, for example, has a different risk profile and requires a different audit focus than a retail chain.

The objective of audit planning is to create an audit plan that can be performed in an effective manner (ISA 300.4) and that makes it possible to reach an audit opinion by the agreed date and to deliver the audit report on time. The audit plan includes the nature, timing, and extent of audit procedures (ISA 300.A12).

The audit planning stage is divided into the following steps (Arens et al., 2017):

- Analysis of audit object
- Risk assessment
- Determination of materiality
- Audit plan

2.1.1 Analysis of audit object

Every company exists to pursue a defined purpose. To achieve this goal, each company interacts with various stakeholders such as customers, employees, suppliers, and shareholders. These interactions create a complex network, which entails various risks (Knechel, 2007). Each company and each industry works differently and, consequently, has different risks. Whereas an international mining company might be more exposed to international developments (e.g., exchange rate fluctuations, global commodity prices), these factors will not be as relevant for a local dairy company, which sells its products only within the region. The auditor is required by ISA 315.A1 to obtain a holistic "understanding of the entity and its environment". Consequently, the auditor must achieve a comprehensive understanding of the audit client and its surrounding network (e.g., suppliers, clients) in order to assess risks correctly.

To gain the necessary knowledge about the client, Arens et al. (2017) further divide this task into the following subtasks:

- Industry and external environment
- Business operations and processes
- Management and governance
- Objectives and strategies
- Business performance

2.1.1.1 *Industry and external environment*

Each industry has its specific risks, which influence the audit to be carried out, and deals with different accounting specifics, which the auditor should consider in the audit program. For example, a construction company has a different accounting focus and revenue streams than a pharmaceutical company. For the former, the realization of sales combined with the percentage of completion will be a central issue. For the latter, research and development play an important role.

Concerning the client's external environment, the auditor must know the current economic situation, the extent of competition, and the regulatory requirements. Local and global economic interrelations impact every client in some way (Arens et al., 2017). Capital markets operate on a global scale, and money can be raised on stock exchanges around the world with international investors participating (Knechel, 2007).

2.1.1.2 *Business operations and processes*

The auditor should understand the client's products, services, customers, suppliers, investors, and related parties (Arens et al., 2017). Each industry operates differently, and every company develops its processes to deal with its environment. The processes and the corresponding internal controls are a main focus for the auditor during the risk assessment and internal controls audit. These processes are covered in more detail in 2.1.2 Determination of materiality and 2.2.1 Internal control audit.

2.1.1.3 *Management and governance*

The investors and owners entrust management with conducting the entity's operations. As a result, the management principles and style – the "tone at the top" – play an essential role in auditing, especially regarding the audit plan and risk assessment. Besides conducting interviews with management, the minutes of board meetings and the audit committee should be reviewed, so that the auditor is informed about leadership decisions, their implementation, and consequences (Arens et al., 2017). The auditor needs to become familiar with the corporate governance practices in place and how they are applied in day to day operations (ISA 260.A14).

2.1.1.4 *Objectives and strategies*

To obtain an overall picture of the audit object and its associated risks, the auditor must inform himself[4] about the defined corporate strategies and objectives. They reflect the goals the client wants to achieve in interaction with the external environment, for example, targeting customers and seizing market opportunities. Those defined goals influence management and employee

---

[4] The male form chosen in this dissertation always refers to both female and male.

behavior as well as the definition of relevant key performance indicators (KPIs) (Knechel, 2007).

2.1.1.5 *Business performance*

As part of the audit planning, the auditor is required to perform analytical procedures to obtain a better understanding of the entity's activities and possible areas of risk (EXPERTsuisse, 2015). As a basis, it is advisable to analyze the most recent available financial statements. A multi-year comparison of the company or average values from industry/competitors can be used as a benchmark for analysis (Arens et al., 2017). By performing this task, the auditor becomes familiar with a company’s capital structure and its financing activities, which are essential indicators for assessing the going concern assumption (Hayes, 2005).

Additionally, the chosen accounting standard plays a vital role for the auditor, since this is the target against which the auditor gives assurance in all material aspects (Hayes, 2005). Therefore, it is recommended that the client’s accounting policies, such as revenue recognition, be reviewed, to make sure they are aligned with the selected accounting standard.

Based on the insights gained about the client and its environment, the auditor can carry on with the risk assessment for the audit engagement.

2.1.2 Risk assessment

The auditor faces the risk of erroneously expressing an unmodified opinion when the financial statements contain material misstatements (audit risk) (Louwers, Blay, Sinason, Strawser, & Thibodeau, 2018). In order to minimize this risk to an appropriate level, the auditor performs a risk assessment (Marten et al., 2015).

Audit risk can be summarized as a qualitative equation with the following factors:

Audit Risk = Inherent Risk x Control Risk x Detection Risk

2.1.2.1 *Inherent risk*

Inherent risk is defined as the probability that material errors occur in the financial statements, assuming there are no internal controls implemented (Arens et al., 2017). In assessing the inherent risk, the auditor considers several factors including the client's industry, findings from previous audits, related party transactions, and complex or unusual transactions.

2.1.2.2 *Internal control and control risk*

The implementation of internal controls mitigates the inherent risk. ISA 315.4 defines internal control as "the process designed, implemented and maintained by those responsible with governance, management and other personnel to provide reasonable assurance about the achievement of an entity's objectives with regard to reliability of financial reporting, effectiveness, and efficiency of operations, and compliance with applicable laws and regulations". Hence, the auditor first needs to gain a rough idea about the internal controls. After that, he assesses the control risk for the audit engagement.

Control risk is defined by Marten et al. (2015, p. 232) as the probability that "material misstatements will not be prevented or detected in a timely manner by internal controls". The more effectively designed the internal controls that are in place, and the more they are systematically and rigorously followed, the fewer issues the auditor can expect to find during the audit engagement (Knechel, 2007). However, there is also a trade-off from a business perspective, because internal controls can impede the process flow, and more controls are not necessarily better for business operation.

2.1.2.3 *Detection risk and audit risk*

The detection risk reflects the probability that the auditor will not detect material misstatements (Kartscher et al., 2013). In contrast to inherent risk and control risk, neither of which can be influenced by the auditor, detection risk arises after an acceptable audit risk has been defined (Marten et al., 2015).

When determining the audit risk, the inherent risk and control risk are given at the beginning of an engagement while detection risk is the residual. The defined detection risk has a direct impact on the audit plan, audit procedures, materiality, and the amount of audit evidence required.

Using the information gathered so far, the auditor determines an acceptable level of audit risk. Quantifying any audit risk is difficult in practice, that is why, three qualitative assessment levels – low, medium, and high – are used (Kartscher et al., 2013).

2.1.2.4 *Fraud*

Even though those charged with governance bear the primary responsibility for the prevention and detection of fraud, the auditor is required by ISA 240 to carry out specific audit procedures in this regard. The auditor has to "identify and assess the risks of material misstatements of the financial statements due to fraud" (ISA 240.10.a) as well as obtain "sufficient appropriate audit evidence regarding the assessed risks of material misstatement due to fraud, through designing and implementing appropriate responses" (ISA 240.10.b). ISA 240 presumes that revenue recognition and management override of controls hold significant risk of fraud. Thus, the standard setter requires the auditor to specifically address these two areas. For revenue recognition, the auditor needs to evaluate which types of revenue, revenue transactions, or assertions are potentially concerned, and address them accordingly (ISA 240.26). Journal entry testing is obligatory because of the potential dangers associated with management override of controls (ISA 240.32). The assessment of the concerned risk takes place during the audit planning stage. However, the actual journal entry testing is conducted during the audit performance stage.

A distinction is made between internal and external fraud. Someone outside of the company commits external fraud, whereas internal fraud results from actions by insiders (e.g., an employee or management).

Internal fraud is further split into the categories of financial statement fraud and transaction fraud (Jans, van der Werf, Lybaert, & Vanhoof, 2011). Financial statement fraud is “the intentional misstatement of certain financial values to enhance the appearance of profitability” (Bologna & Lindquist, 1995). This type of fraud is also known as management fraud, since it usually involves a management override. Transactional fraud can be committed by anyone within the company and is defined as stealing or embezzling organizational assets (Jans et al., 2011). When it comes to occupational fraud (employees stealing from the employer), auditors identify only 3% of the cases, based on a study published by the Association of Certified Fraud Examiners (ACFE) (2019).

Detecting fraud is a demanding task for the auditor, especially when the fraud is well thought through and carefully planned. In Switzerland, the monetary damage of an identified fraud case is on average around CHF 9.5 million (Mautone, Häni, & Völker-Alber, 2018).

2.1.3 Determination of materiality

The ISA standard defines misstatements to be material “if they, individually or in the aggregate, could reasonably be expected to influence the economic decisions of users taken on the basis of the financial statements” (ISA 320.2). To determine materiality the professional judgment of the auditor is required. The determination of materiality plays an essential role in the audit process as it is the auditor's responsibility to confirm that the audited financial statements are free from material misstatements (EXPERTsuisse, 2015).

The determination of materiality is a balancing act. It requires the auditor to have the users of the financial statements in mind and to apply all his knowledge and experience in making a "professional judgment" (Kartscher et al., 2013).

Furthermore, materiality also depends on the performed risk assessment. The threshold for judging that a material misstatement has occurred is lower when the audit risk is higher. Conversely, the lower the audit risk, the higher the materiality threshold. The materiality threshold may not be exceeded by the aggregate of all errors detected or in case of errors in samples, the extrapolation of errors, as otherwise, the financial statements will be materially misstated. This is important for the auditor, because he has to confirm that the financial statements are correct in all material matters.

In the course of the audit, various materiality levels are determined. Several types of materiality and their respective uses are briefly explained in the following sections.

2.1.3.1 *Overall materiality*

As a first step, the auditor sets the materiality threshold for the entire financial statements. This value represents the maximum amount of misstatements that could exist in the financial statements without affecting the financial statement user's decisions. This limit may be adjusted during the audit and is used to determine an audit strategy (ISA 320.10).

2.1.3.2 *Specific materiality*

Specific materiality thresholds are established for certain types of transactions, account balances, or disclosures (ISA 320.10). The determination of the specific materiality is again the responsibility of the auditor exercising professional judgment. He needs to have the financial statement users in mind and the potential impact of a specific item on their economic decisions.

2.1.3.3 *Performance materiality and non-access threshold*

Performance materiality is used to reduce the risk of material misstatements to an appropriately low level so that aggregated uncorrected findings and undetected misstatements do not exceed the overall materiality (ISA 320.9).

The materiality concept does not just apply to a maximum amount. It also includes a lower amount to make sure the auditor's focus is not distracted by small errors. Findings below this limit, called the non-access threshold, are irrelevant (EXPERTsuisse, 2015).

2.1.4 Audit plan

As the final step in the audit planning stage, the auditor prepares an audit plan, including an audit strategy, which serves as a guideline for the preparation of the audit program, and the audit program itself.

The auditor is obligated by ISA 260 to communicate to "those charged with governance" (e.g., the executive board) the planned scope and timing of the audit. Furthermore, it is required by the standard that the auditor establishes a communication process that includes "the form, timing and expected general content of communications" (ISA 260.18). The communication process also includes communication about the auditor's responsibilities in relation to the financial statement audit.

2.1.4.1 *Audit strategy*

Based on the information obtained about the audit client and their business, an overall audit strategy is developed. The audit strategy includes the scope, timing, and direction of the audit (ISA 300.7). In addition, it consists of the characteristics of the engagement that define the scope, considers significant results gained during the audit planning stage, and mentions the nature, timing, and extent of resources to perform the engagement (ISA 300.8).

2.1.4.2 *Audit program*

As the final step in the audit planning stage, the auditor draws up a detailed audit program, which operationalizes the overall audit strategy (Porter, Simon, & Hatherly, 2011). When planning the audit, the auditor follows a risk-based approach to focus on the relevant topics, and thus makes sure to lower the audit risk to an acceptable level (EXPERTsuisse, 2015). The audit program includes the nature, timing, and extent of audit procedures, as well as the supervision of the engagement team, and review of their work (ISA 300.11). While performing the audit, the auditor may need to adjust the audit strategy and audit program if necessary (ISA 300.A10).

2.2 Audit performance

The audit plan is carried out during the execution of the audit. This chapter explains the audit techniques carried out during this stage. In practice, the audit performance stage is split in two: the preliminary (before the balance sheet date) and the main audit (after the balance sheet date).

This stage aims to collect sufficient and appropriate audit evidence to form an opinion about the financial statements. ”Sufficient” refers to the quantity of audit evidence, which is “affected by the auditor’s assessment of the risks of material misstatement and also by the quality of such audit evidence” (ISA 500.5.e). ”Appropriate” is a measure of quality, which is expressed in terms of relevance and reliability “in providing support for the conclusions on which the auditor’s opinion is based” (ISA 500.5.b).

The “audit performance” stage contains the internal control audit and substantive procedures (Marten et al., 2015), each of which are split further into two subgroups. The internal control audit consists of the test of design and test of controls, while the substantive procedures involve substantive analytical procedures and test of details.

Before proceeding to consider the procedures of the audit performance, it is noted that for two of them, the test of controls (see 2.2.1.2) and the test of details (see 2.2.2.2), the auditor can choose whether to audit the whole population or a sample of it. Sampling methods have been integrated into the audit profession for a long time, because it is not feasible to examine every transaction and account balance (Porter et al., 2011). Moreover, sampling is efficient and cost-effective compared with testing the full population, which is often practically impossible given the amount of information. The auditor is required to check if the financial statements are represented according to the accounting framework in all material matters. However, auditing just a sample entails some risk of arriving at a wrong impression. An auditor might, for example, conclude that internal controls are more effective than they actually are (ISA 530.5).

2.2.1 Internal control audit

The term "internal control audit" does not occur explicitly within ISA, but is used as an overarching term in the literature (e.g., Marten et al., 2015). As the name indicates, the focus of this audit substage is the client's internal control system, of which the auditor will have already gained a rough idea during the audit planning stage.

The internal control audit is split up in two parts. First, the auditor gets an understanding of the existing internal controls and evaluates how they are designed ("design effectiveness"). Then, the test of controls in terms of effective functionality takes place ("operating effectiveness"). This recognizes the fact that even for well-designed internal controls, it is not guaranteed that they are actually applied and effectively implemented (Marten et al., 2015). Conversely, there is little use in assessing the operating effectiveness of the controls if the design of the controls is poor (ISA 315.A66).

2.2.1.1 *Test of design*

For this step, the auditor tests the "design effectiveness". Therefore, understanding the nature and extent of each relevant control is the main focus. The auditor is responsible for understanding and assessing the appropriateness of the key internal controls, which are directly related to the financial statement outcome of the audit client (ISA 315.22).

The Committee of Sponsoring Organizations of the Treadway Commission (COSO) developed an internal control framework, which has become the most widely accepted standard for defining and assessing internal controls (Moeller, 2011). This framework has become a part of best practices for auditing internal controls and its application is expected by ISA (Marten et al., 2015).

COSO developed a three-dimensional model to describe an internal control system that illustrates relevant interactions between various factors.

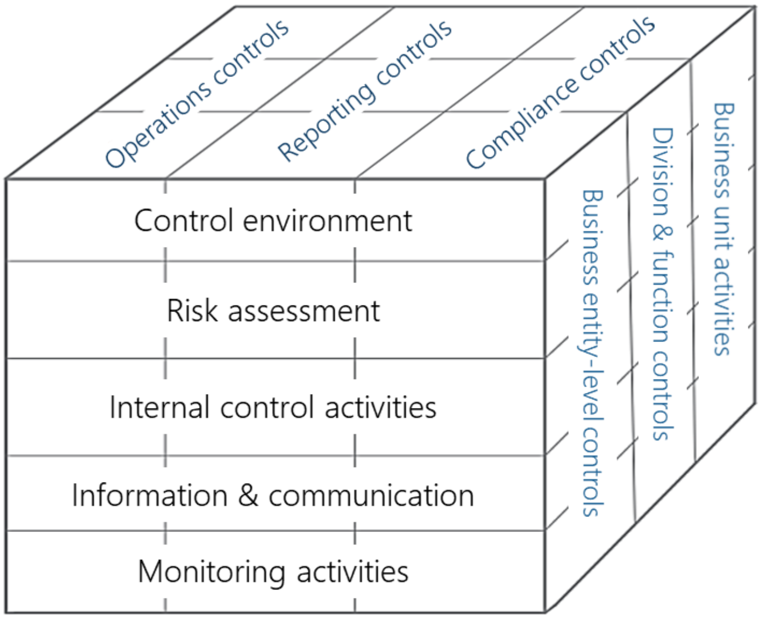

Figure 3 Internal control framework retrieved from the Committee of Sponsoring Organizations of the Treadway Commission (2013)

The front of this cube depicts the five components, which apply across all entity levels and to each objective (operations, reporting, compliance). In order to assess the internal controls of the audit client as effective, all principles shown must be present and functioning (Arens et al., 2017).

In obtaining an understanding of the control environment, the auditor is instructed by ISA 315 to evaluate not only the appropriateness of the internal control components, but also the honesty and ethical behavior of the management ("tone at the top"). "A poor control environment bears the risk that internal controls are not performed or performed just as a matter of form," according to Marten et al. (2015, p. 308). A positive control environment is seen as a necessity, but not in itself sufficient to attest "operating effectiveness".

Each company faces specific risks when trying to achieve the firm's defined objectives. The auditor needs to understand those risks in order to assess their impact on the financial statements via potential occurrences of material misstatement (Knechel, 2007).

Various control activities are installed in workflows to mitigate the identified risks (Moeller, 2013). For example, every purchase order above CHF 1’000 may need approval by the relevant supervisor in a given organization. Such controls can be performed manually or built directly into the accounting system/Enterprise-resource-planning-system (ERP-system). It is crucial for the effectiveness of internal controls that they are documented and communicated across the company (Knechel, 2007).

Lastly, the control activities that have been developed and implemented need to be continuously monitored and, where necessary, adjusted. The auditor must check if the internal controls are monitored and arising deficiencies are handled appropriately (ISA 315.22ff).

In order to obtain the necessary understanding of the internal control design, the auditor performs various audit procedures (EXPERTsuisse, 2015; ISA 315.A67). One audit procedure is the examination of the internal control documentation as the previously presented COSO framework requires the creation and maintenance of many documents and records (Arens et al., 2017). By examining those documents, the auditor can evaluate if the control activities are effectively designed. In case the auditor conducts inquiries, statements from the client can be compared with internal control documentation, and red flags are raised where there are deviations.

When conducting a walk-through test, the auditor selects a business transaction and follows it through the entire accounting process (Arens et al., 2017). For example, purchase transactions can be traced from the purchase order to the invoice, control of incoming goods and inventory accounting, through the general ledger and accounts payable sub-ledger to the payment of the invoice (EXPERTsuisse, 2015). Performing a walk-through test can be time-intensive.

In addition, the auditor can create flow charts to visualize the different processes, which helps to get a comprehensive overview of how the various processes impact the accounting process and the financial statement.

Finally, by conducting inquiries and observing employees, the auditor gets a better understanding of how controls are carried out and implemented. It should be noted that inquiries, in which management or client employees explain what control activities are performed and how they are carried out, are not by themselves sufficient (ISA 315.A67). They can, however, help the auditor to evaluate whether management and employees fully understand their respective duties (Arens et al., 2017).

Once the information is gathered, the auditor can evaluate the design of the internal controls.

2.2.1.2 *Test of controls*

After evaluating the design, the auditor carries out the test of controls to check the operating effectiveness. ISA 330.4 defines tests of controls as “an audit procedure designed to evaluate the operating effectiveness of controls in preventing, or detecting and correcting, material misstatements”. If the auditor wants to rely on the operating effectiveness of the controls in the course of the audit, he is obliged by the standard to obtain sufficient and appropriate evidence for their operating effectiveness (ISA 330.8). As Arens et al. (2017, p. 417) state, “Each key control […] that the auditor intends to rely on to support a control risk below the maximum must be supported by sufficient tests of controls”. This could be achieved by selecting a sample to check the operating effectiveness of the internal controls and their continuous application. For example, the auditor may perform a three-way match, meaning that a purchase order, invoice, and delivery note are compared.

During the planning stage, the auditor will have already made a preliminary judgment about the entity's control risk. However, this judgment will still need to be validated with the test of controls (Arens et al., 2017). To what extent the test of controls is applied depends on the assessed control risk and the judgment of the auditor (Hayes, 2005).

The following audit procedures (examination of records/documents, functionality checks, reperformance, IT-controls, observation and inquiry), which are depicted in the audit process, can be applied for tests of controls, according to EXPERTsuisse (2015). The performance of designed control activities leaves a trail of documentation, records, and reports, electronically and/or on paper. These documents allow the auditor to assess the performed internal controls in terms of operating effectiveness. An absence of documentation indicates that the designed control was not performed.

EXPERTsuisse (2015) also suggests conducting functionality checks on the operating effectiveness of internal controls. For example, an audit client may have designed an internal control where inventory write-offs exceeding CHF 10’000 need written approval from the CFO. In case of such write-offs, the auditor could check if the written approval is available.

In some cases, documents and records exist, whose content is insufficient to evaluate the effectiveness of the internal control. To cover such cases, the auditor can reperform procedures. Arens et al. (2017) give an example of prices on sales invoices, which are obtained from the master price list. Just seeing the price on the invoice does not give the auditor enough evidence that the prices were actually taken from the master price list. Therefore, the auditor can reperform the procedure by checking the indicated prices with the authorized price list. It can be assumed that the procedure was operating as intended if no deviations are identified.

Another potential source for identifying the operational ineffectiveness of controls can be reports from internal audits (EXPERTsuisse, 2015).

To rely on general IT-controls, the auditor usually involves an IT-auditor, who audits those controls and has the expert knowledge to evaluate their appropriateness. In larger companies, in particular, business processes are reflected fully or in part in ERP-systems (Kartscher et al., 2013). Internal controls are programmed into these systems. Unless those controls have been amended during the period under audit, the auditor can expect that the controls function consistently and therefore perform a “test of one” for those controls (ISA 330.A29).

One possible audit procedure is the observation and questioning of employees. The auditor can either question people who perform the control procedures or those that monitor the controls, such as the IT-administrator who hands out access rights to specific applications (Hayes, 2005). Observations of procedures being performed are especially useful for processes that leave no audit trail in the system. However, this may be more appropriate for the test of design than for the test of controls. As mentioned above, inquiry alone is not sufficient audit evidence (ISA 315.A67).

Some audit procedures are also used for the test of design. However, the focus for test of controls is the actual execution of the designed controls (Arens et al., 2017). During the test of design, the auditor gets an understanding of all internal controls, whereas, during the tests of controls, specific controls are assessed for effective operation.

At the end of the test of controls, the auditor needs to re-evaluate his previous judgment about the effectiveness of the internal controls. In cases where they were proven to be effective during the audit, a reduced control risk is supported by the collected audit evidence (Hayes, 2005). In case of significant deficiencies being identified, the auditor is obliged by ISA 265 to communicate the findings in writing and to adjust the preliminary judgment made during the assessment of the control risk. In this context, deficiencies exist when "a control is designed, implemented or operated in such a way that it is unable to prevent, or detect and correct, misstatements in the financial statements on a timely basis" (ISA 265.6.a).

ISA 315.A46 recognizes the inherent limitations of the internal control system and states that reasonable assurance is only provided if the effectiveness is proven. To provide evidence of this, the performance of substantive procedures is necessary.

2.2.2 Substantive procedures

Substantive procedures focus on testing the disclosures made by the client within the financial statements (Marten et al., 2015). They are defined by ISA 330.4 as audit procedures "designed to detect material misstatements at the assertion level". In the course of the preparation of the financial statements, the audit client's management made various assertions regarding "recognition, measurement, presentation and disclosure of the various elements of financial statements and related disclosures" (ISA 315.A110). With those assertions, the management claims that the financial statements were prepared following the objectives of the applied accounting standard (Marten et al., 2015). In the case of IFRS, the objective would be that the financial statements are presented fairly (International Accounting Standard 1), which is known as the true and fair view principle. Therefore, the auditor has to check if the management assertions, also known as financial statements assertions (Marten et al., 2015), meet the respective accounting standard's objectives and are free from material misstatements.

ISA 315.25b classifies the assertions into three categories: classes of transactions and events, account balances and presentation and disclosure. Each category includes various assertions, which aim to help the auditor to identify potential misstatements. Table 1 (next page) lists all assertions for each category, in accordance with ISA 315.

Assertions about classes of transactions and events	
Occurrence	Transactions and events that have been recorded have occurred and pertain to the entity.
Completeness	All transactions and events that should have been recorded have been recorded.
Accuracy	Amounts and other data relating to recorded transactions and events have been recorded appropriately.
Cut-off	Transactions and events have been recorded in the correct accounting period.
Classification	Transactions and events have been recorded in the proper accounts.
Assertions about account balances at the period end	
Existence	Assets, liabilities, and equity interests exist.
Rights and obligations	The entity holds or controls the rights to assets, and liabilities are the obligations of the entity.
Completeness	All assets, liabilities and equity interests that should have been recorded have been recorded.
Valuation and allocation	Assets, liabilities, and equity interests are included in the financial statements at appropriate amounts and any resulting valuation or allocation adjustments are appropriately recorded.
Assertions about presentation and disclosure	
Occurrence and rights and obligations	Disclosed events, transactions, and other matters have occurred and pertain to the entity.
Completeness	All disclosures that should have been included in the financial statements have been included.
Classification and understandability	Financial information is appropriately presented and described, and disclosures are clearly expressed.
Accuracy and valuation	Financial and other information are disclosed fairly and at appropriate amounts.

Table 1 Assertions according to ISA 315.A111

An example of an account balance assertion can be the completeness of all bank loans. Positive confirmation requests (ISA 505.6b) can be sent out to the banks by the audit client to test this assertion. Such bank requests contain either a loan balance on the due date, which needs to be confirmed, or a request to provide the loan balance on a specific date. The confirming party needs to respond to the auditor directly.

ISA 330.18 requires that the auditor designs and performs “substantive procedures for each material class of transactions, account balance, and disclosure”. The substantive procedures are subdivided into substantive analytical procedures and tests of details. Which of these audit procedures is used, or whether a mix of both is appropriate, depends on the auditor's judgment. The performance of substantive procedures is obligatory for the subjective risk assessment (ISA 330.A42).

2.2.2.1 *Substantive analytical procedures*[5]

The term “analytical procedure” is defined by ISA 520.4 as the evaluation “of financial information through analysis of plausible relationships among both financial and non-financial data”. In other words, analytical procedures assess the consistency and economic plausibility of specific values or sets of business transactions (Marten et al., 2015).

Analytical procedures are performed during each audit stage, but with a different purpose in each stage. During the planning stage, analytical procedures enhance the auditor’s understanding of the audit client and its business. During the audit performance, analytical procedures serve to obtain evidence about assertions, and are called substantive analytical procedures.

---

[5] For brevity this procedure is shown as “analytical procedures” in the audit process.

Finally, during audit completion, analytical procedures are used to assist in forming an overall conclusion.

Before performing substantive analytical procedures, the auditor must evaluate if they are suitable for detecting material misstatements (ISA 520.A6). For example, substantive analytical procedures can be suitable for large volumes of transactions (ISA 330.A44), but do not alone provide sufficient appropriate evidence (ISA 330.A10). For small volumes of transactions, a test of detail might be more appropriate.

When performing substantive analytical procedures, the auditor needs to set an expected value as well as a tolerance range, which allows for some deviation (Marten et al., 2015). For example, if sales significantly increased, a rise in the cost of goods sold can be expected. If material differences appear between the auditor's expectations and the results of the analytical procedures performed, additional evidence needs to be gathered to identify potential misstatements (Knechel, 2007).

In cases where the auditor performs the analysis based on information provided by the client, he needs to ensure first that the information is complete and accurate. Otherwise, the information cannot be relied on. For example, if the auditor receives a list of all outstanding receivables, he should check that the total sum is reconcilable with the general ledger amount for receivables.

Various analytical procedures can be used in an audit, depending on the defined audit objective (EXPERTsuisse, 2015). One simple procedure is the comparison of the current and previous year or budget versus the current year. Significant changes which cannot be explained by the client's environmental factors can be a first indicator of errors in the financial statement and suggest a need for further investigation (Marten et al., 2015).

When comparing the current year with former years, material changes of accounting methods during both periods must be considered, and correcting adjustments may need to be made to enable comparability. When performing an analysis of budget versus actual, the auditor should consider that the audit client might have incentives to meet the targets in the budget. Therefore, the auditor must critically question the budget figures.

Another example is the calculation and analysis of KPIs to help the auditor analyze the economic and financial situation of the audit client over time (EXPERTsuisse, 2015). This extends beyond financial KPIs such as liquidity or profitability ratios, to include productivity and utilization ratios. When using KPIs over time, experience reveals practical issues that emerge. A lack of consistency can arise because companies evolve over time, and their financial statements change (e.g., business acquisition or merger). When such changes have occurred, significant effort may be required to arrive at a meaningful set of data from which KPIs can be measured.

Analyzing the financial statements helps the auditor gain an overview of the events in the period under audit. This was first done during the audit planning stage to help the auditor gain a better understanding of the audit object as well as for the initial risk assessment. EXPERTsuisse (2015) suggests the following procedures when analyzing the financial statements during the audit performance: As a first step, the auditor should look at the structure of the financial statement and determine if it is following the accounting standard. Looking at the balance sheet as well as the profit and loss statement can also reveal valuable insights regarding capital structure or going concern issues.

Certain business transactions influence the balance sheet and the profit and loss statement, as well as the notes. For example, sales influence revenues, bank/cash/accounts receivables, and disclosure. Because such interrelations are a given fact, the auditor can perform plausibility checks. These include simple measures such as whether various costs as a percentage of revenue stay stable, and extend to far more complex measures that involve combining financial statements with additional KPIs to calculate, for example, unit economics.

Using trend analysis as analytical procedures can indicate the client's likely business development in the future based on the performance in the past. It is recommended that non-periodic, non-operating, and extraordinary results are corrected to gain better results for the trend analysis (EXPERTsuisse, 2015).

Also, benchmarking can be conducted to compare the audit client with its competitors (e.g., performance) or to make comparisons within the company itself (e.g., between different production facilities). Through such benchmarking, strengths and weaknesses can easily be identified. Within the context of an audit, the goal is not to make suggestions for improvement, as may be the case in a consulting engagement, but rather to identify areas of potential risk (Marten et al., 2015).

2.2.2.2 *Test of details*

Whereas substantive analytical procedures are used to gain a high-level impression and check the plausibility of the information, tests of details are performed to gain detailed insights. They are especially useful for testing assertions.

When conducting a test of details, the auditor must decide whether to audit the entire population or just a sample (ISA 530.5). The scope and selection of audit elements play an essential role in the test of details. ISA is aware that auditing just a sample bears some risk, particularly when it comes to drawing a conclusion about the population based on the sample selection (ISA 530.5c). The sample size decision depends on the level of sample risk the auditor is willing to accept (ISA 530.A10).

For selecting items for testing, the standard indicates methods such as random selection, systematic selection, and monetary unit sampling (ISA 530.A13; ISA 530 Appendix 4). As the name “random selection” implies, a random sample is selected with the help of a random number generator. Each transaction within the population has an equal chance of being drawn. In a systematic selection, a specific rule is used to pick transactions; for example, every 50th transaction is selected for review. Monetary unit sampling is a value-weighted selection, where individual monetary units make up the population (ISA 530.5).

If misstatements or errors are identified in the tested sample, the auditor has to extrapolate the number of errors in the sample to the whole population. In cases where the extrapolated amount is above the defined materiality level, adjustments need to be made.

The following audit techniques for a test of details are possible, and are explained in more detail below: examination of records/documents; recalculation; reconciliation; critical review; and observation, inquiry, and inspection (EXPERTsuisse, 2015).

Performing an examination of records and documents as a test of details can have two focuses.

First, each journal entry should be based on a source document, such as an invoice or receipt. Therefore, the auditor inspects records and documents to check if they were completely and accurately captured in the accounting system (e.g., amount, currency, consistency of information on the document and in the system) (EXPERTsuisse, 2015). For this audit procedure, the auditor needs to consider both externally produced receipts and internal ones. For automatically entered journal entries, test of controls procedures might be more appropriate.

Second, examination of records and documents can be also used to gain further insights about the audit object, especially in situations where facts might not be directly reflected in the financial statements or accounting system (e.g., guarantee declarations, litigation risk) (EXPERTsuisse, 2015). Other sources might be meeting minutes of the board of directors or audit committee, correspondence with lawyers, and internal audit reports, to name a few.

Another possible audit procedure is recalculation, which is used to uncover calculation errors. This consists of checking mathematical accuracy manually or electronically (ISA 500.A19). Due to the increased use of IT-systems, recalculation is no longer used as frequently as it was when more manual calculation was involved in creating financial statements (EXPERTsuisse, 2015). Therefore, the auditor today is more focused on whether the system calculates accurately, which is part of checking the IT-controls.

With the help of cross-referencing and reconciliation, the auditor can cross-check the client's numbers in various documents for consistency with numbers in the ledgers. For example, the numbers in the financial statement notes can be cross-referenced with the numbers from the general ledger or sub-ledger.

The critical review of documents can give the auditor a better understanding. For example, the purchasing contracts of new companies or management's compensation packages might be subjected to a critical review.

For the test of details, observation, inquiry, and inspection are again possible audit techniques. A typical example of a test of details observation is attendance at the inventory count. Performing follow-up inquiries based on identified deficiencies during substantive analytical procedures may be another suitable action.

2.3 Audit completion

The audit is complete once the audit report is handed over to the client. The final stage of an audit focuses on reviewing the collected audit evidence, subsequently forming an opinion, and lastly, writing the audit report.

2.3.1 Going concern

When preparing the financial statements, management uses the going concern assumption unless liquidation plans exist. Based on ISA 570, the auditor is obliged to either validate this assumption or to determine if the business is likely to fail in the foreseeable future. In cases where the auditor has material uncertainty about meeting the going concern assumption based on the collected evidence, the valuation of the financial statement needs to be reconsidered, and certain actions may be required as outlined by ISA 570 and applicable local regulation.

2.3.2 Audit opinion

As part of the judgment process, the auditor reviews the audit documentation and addresses open issues. Reviewing the audit documentation as well as the relevant and reliable audit evidence that has been collected helps the auditor to form an opinion about the audited financial statements. It also ensures that all relevant audit results are documented, and that open issues and missing documents are recorded (EXPERTsuisse, 2015).

Once the auditor has reviewed the documentation and reached a judgment, he prepares the final report. If the auditor concludes that "the financial statements are prepared, in all material respects, in accordance with the applicable financial reporting framework" (ISA 700.7.c), an unmodified opinion will be expressed in the audit report. Otherwise, if the auditor concludes that the financial statements are not prepared in all material matters following the applicable financial reporting framework, a modified opinion is issued. ISA 705.2 names three types of modified opinions: qualified opinion, adverse opinion, and disclaimer of opinion. Table 2 lists the ISA regulations applicable to each modified opinion.

Modified Opinion	ISA Regulation
Qualified opinion	ISA 705.7: “(a) The auditor, having obtained sufficient appropriate audit evidence, concludes that misstatements, individually or in the aggregate, are material, but not pervasive, to the financial statements; or (b) The auditor is unable to obtain sufficient appropriate audit evidence on which to base the opinion, but the auditor concludes that the possible effects on the financial statements of undetected misstatements, if any, could be material but not pervasive.”
Adverse opinion	ISA 705.8: “The auditor shall express an adverse opinion when the auditor, having obtained sufficient appropriate audit evidence, concludes that misstatements, individually or in the aggregate, are both material and pervasive to the financial statements.”
Disclaimer of opinion	ISA 705.9 "The auditor shall disclaim an opinion when the auditor is unable to obtain sufficient appropriate audit evidence on which to base the opinion, and the auditor concludes that the possible effects on the financial statements of undetected misstatements, if any, could be both material and pervasive.”

Table 2 Modified opinions based on ISA 705

2.3.3 Audit report

In accordance with ISA 260, the auditor communicates the conclusion to those charged with governance (e.g., board of directors) of the audited firm through the audit report. The recipients are defined as people or organization(s) "with responsibility for overseeing the strategic direction of the entity and obligations related to the accountability of the entity" (ISA 260.10.a).

The auditor informs about significant findings from the audit, including material qualitative aspects of the entity's accounting practices and any substantial difficulties encountered during the audit. If applicable, discussions of significant matters with management need to take place (ISA 260.16) and require documentation (ISA 230). The auditor also has to inform those charged with governance about uncorrected misstatements and the corresponding impact on the financial statements (ISA 260.12). Depending on the local regulations, additional reports might be issued.

2.4 Documentation

Since 2009, the auditor is required by ISA 230 to sufficiently and appropriately document the audit. This serves the purpose of creating evidence that the audit was planned and performed in accordance with ISA and other applicable legal and regulatory requirements (ISA 230.5). The auditor is required to document the performed audit procedures, evidence obtained, and conclusions reached (ISA 230.6.a). The working papers can be stored in electronic or physical form (ISA 230.6.b) and need to be prepared on a timely basis (ISA 230.7).

ISA 230 also gives guidance about the form, content, and extent of the audit documentation. When storing the information in files, a distinction is made between a permanent audit file and a current audit file (Porter et al., 2011). The former stores legal documents and general information about the client, among other things. The latter file structure contains more working documents, such as the current year's audit plan or a list of unresolved issues.

The audit documentation, including documents prepared by the client, is the property of the auditor (Porter et al., 2011). Therefore, he is responsible for its confidentiality, safe custody, integrity, accessibility, and guaranteed retrievability (ISA 230.46ff).

2.5 Quality control

As a reaction to the financial crisis in 2008, IAASB published two quality standards, ISA 220 and ISQC 1, to ensure a certain level of audit quality (Marten et al., 2015). ISA 220 regulates the quality within an audit engagement, whereas ISQC 1 governs the quality of an entire audit organization.

Quality control during the audit process is a review process providing "an objective evaluation, on or before the date of the auditor's report, of the significant judgments the engagement team made and the conclusions it reached in formulating the auditor's report" (ISA 220.7.b). For guarding audit quality, an "engagement quality control reviewer" is assigned. This is a person within the firm but outside the immediate engagement team, with sufficient and appropriate experience as well as the authority to objectively evaluate significant judgments and conclusions reached (ISA 220.7.c). Notwithstanding this quality control, the engagement partner remains primarily responsible for audit quality and guarantees that the work is carried out correctly (ISA 220.8).

To ensure both quality standards are obeyed, most countries have regulatory bodies that are responsible for ensuring the proper performance and quality of audit services through enforcement procedures (FAOA, 2019b). If legal violations appear in the course of an enforcement, steps are taken to punish them. In Switzerland, this organization is called the Federal Audit Oversight Authority (FAOA). It is also responsible for licensing individuals and audit firms, as well as issuing rules and standards for the Swiss audit sector.

The theoretical background for the audit process having been discussed, the next chapter focuses on data analytics, with a particular focus on auditing.

3 Data analytics

Data analytics has been around for decades, and started getting more popular around the turn of the millennium (Runkler, 2016). Various forms of data analytics, such as business intelligence, have been implemented in the business world over the past decades (Davenport, 2014). Business intelligence combines various procedures to systematically and frequently analyze internal and external business data (Mertens et al., 2017).

What has changed over the past few years is that the commoditization of data storage and the growth of computing power at affordable prices now provide the prerequisites for creating and storing vast amounts of data (Kelleher & Tierney, 2018). The massive and increasing amount of information digitally available has led to a growing interest in data analytics (Provost & Fawcett, 2013) as well as the various software solutions available to perform them.

Data analytics within this dissertation is defined as "the process of extracting meaningful knowledge from data" (Roiger, 2017, p. 38). It is used as an umbrella term to cover a range of ambiguous terms with overlapping meanings, and which include a broad spectrum of methods that revolve around data – terms such as "data mining", "data science", "text mining", "artificial intelligence", and "machine learning" (Abbott, 2014; Finlay, 2017; Gee, 2014; Han, Kamber, & Pei, 2012; Kelleher & Tierney, 2018; Runkler, 2016).

To describe the process of data analytics within this dissertation and to organize the following subchapters, the Cross-Industry-Standard-Process for Data Mining (CRISP-DM) (Chapman et al., 1999) is used. According to Kelleher and Tierney (2018), it is the most commonly applied process within the field of data analytics.

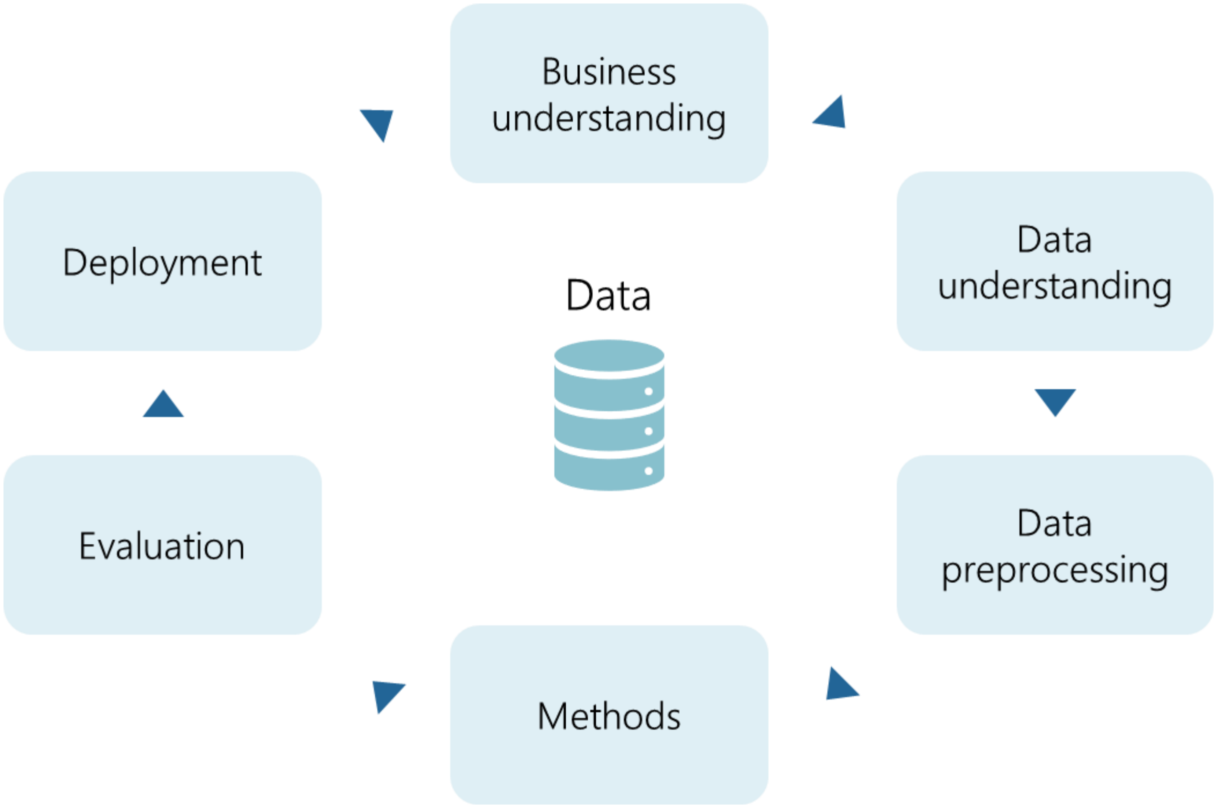

Figure 4 Cross-industry-standard-process for data mining based on Chapman et al. (1999)

The cyclical arrangement of the steps in the CRISP-DM reflects the multiple iterations needed when performing data analytics in practice (Provost & Fawcett, 2013). Data is at the center of the whole process as it is the central subject of every analysis (Kelleher & Tierney, 2018).

3.1 Business understanding

The first step focuses on understanding the business problem or objective of the planned analysis. It is crucial for a successful project to clearly define the problem to be solved (Kelleher & Tierney, 2018). Before starting the data analytics process, evaluation criteria for the final model should be defined (Finlay, 2017). Additionally, the deployment criteria need to be determined to make sure the developed model can be used for the intended application in practice (Abbott, 2014). Provost and Fawcett (2013) recommend systematically breaking down the issue at hand into tasks that can be handled with data analytics methods.

In this stage, it also needs to be ensured that the data and resources required for the analysis are available (Wirth & Hipp, 2000). Already during this stage, an initial assessment of potential methods (e.g., supervised versus unsupervised learning) and their respective advantages and disadvantages should be carried out, since this might trigger more specific requirements for the data.

In times of increasing regulatory requirements around data security and data privacy (e.g., European general data protection regulation), it is necessary to evaluate whether the planned analysis is allowed under the applicable laws (Meyer, 2012). It can be a project killer if the required data is not available for analysis due to regulatory restrictions, or if a violation of law is committed in the process of the analysis.

Starting with a good understanding of the business problem, i.e., formulating the proper hypotheses for analysis, cannot be overstated in terms of importance (Kelleher & Tierney, 2018). Data analytics projects and implementations can fail because they do not have a clear business objective or because there is a gap of understanding between the business owners and the data scientists regarding the proper scope for a data analytics project.

3.2 Data understanding

This stage includes extracting the data from the source system as well as gaining a qualitative understanding of the data at hand. Regarding the application of data analytics to auditing the systems used (accounting software or ERP-systems), as well as the storage facilities (data warehouses), deserve a more detailed consideration.

3.2.1 Sources of data

3.2.1.1 *Accounting software and enterprise resource planning systems*

In an auditing context, the primary interest of the auditor will be to look at data in the accounting software or ERP-system, as the journal entries recorded there provide the basis for the financial statements. A variety of accounting software exists, which leads to the drawback that there is no common data model. The options which are offered by different accounting systems vary. Some provide a system for bookkeeping, whereas others include features such as payroll accounting, invoice writing, or value-added-tax returns. Compared to ERP-systems, accounting software have fewer functionalities.

An ERP-system is an integrated system that supports all operative and management functions of a company (Leimeister, 2015). The system is built in a modular way with function-related modules, e.g., for accounting, procurement, production planning, and human resources. The two leading ERP-providers worldwide are SAP and Oracle (Mertens et al., 2017).

A company can individually adapt an ERP-system to its business processes (Leimeister, 2015). The ERP-system is often tailored to the specific industry and company needs. In some companies, the ERP-system acts as the central nervous system of the company, and as such is often directly connected to various other business-critical data sources. For example, in industrial companies, ERP-systems are frequently directly connected to the manufacturing or processing equipment.

In terms of data architecture, all modules are based on the operative database, which makes it possible to combine business processes across functions (see Figure 5) (Leimeister, 2015). The operative database is the primary storage location for business data, where read-write processes take place. It also contains the most up-to-date version of business data. However, this comes at a cost in terms of the speed of data retrieval, and the operative database can become a bottleneck in terms of system performance (Abts & Mülder, 2017).

3.2.1.2 *Data warehouses*

When considering auditing, which deals mostly with data about the past (Krieger & Drews, 2018), it becomes apparent that the operative database, which contains the latest and most up to date information, may not be the only data-source of interest. Furthermore, specific applications usually do not require all data but rather a particular subset. The solution is to extract a copy or subset of the required data for these applications from the ERP-system's operative database into a data warehouse (Vaisman & Zimányi, 2014).

Transferring data from the operative database to a data warehouse is an ongoing process that happens according to a fixed routine (e.g., monthly, weekly, daily, hourly) (Roiger, 2017). A data warehouse's function is to store data and provide read-only access, to ensure that no changes are made to the data (Roiger, 2017). From there, the data can be used for further analysis (e.g., business intelligence) (Mertens et al., 2017). To ensure that robust results can be derived from the data analysis, adequate data needs to be extracted from the operative database into the data warehouse. Since a data warehouse functions as a repository (see Figure 5), other data, such as data from external sources, can be added (Han et al., 2012).

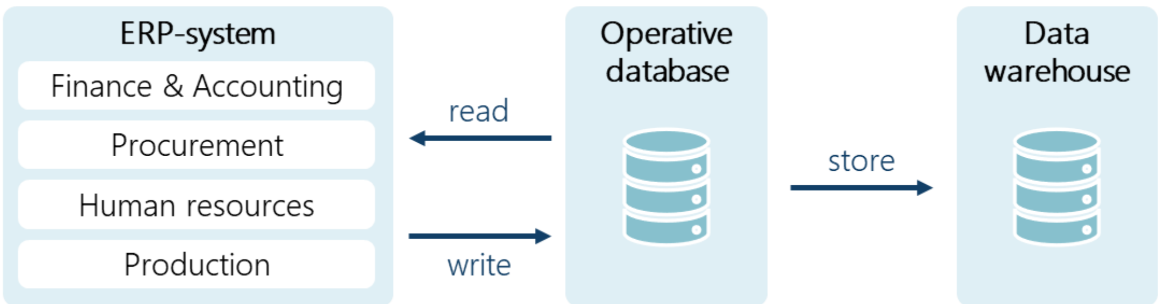

Figure 5 Illustrative structure of ERP-system and data storage based on Abts and Mülder (2017)

In practice, a first hurdle for the auditor is that ERP-systems are not programmed in a common accounting language. In the example of Bönner, Riedl, and Wenig (2011), using accounting terms within a data export request to the IT-department can create confusion when the same terms are not found in the system. Depending on the request, several data export tables may need to be combined to provide the data required to fulfill the request. The auditor knows what kind of data is needed for the analysis, but for some requests it is still the IT department which selects the tables and fields, and performs the data export. Extracting a useful dataset therefore requires cooperation between auditors, who have the accounting knowledge, and the IT-responsible counterparty, who can explain how and where data is stored in the system (Bönner et al., 2011). The auditor then needs to check the completeness of the information provided and reconcile it with other information obtained (e.g., balance sheet).

3.2.2 Types of data

A distinction is made between structured and unstructured data. In essence, structured data "can be stored in a table, and every instance in the table has the same structure (i.e., set of attributes)" (Kelleher & Tierney, 2018, p. 48). There is consistency in a column in terms of data types, with each entry having the same attribute. An example is the general ledger of a company, with each journal entry having the same structure (e.g., debit and credit account, date, amount). Data exported from a source system such as SAP or, more generally, any data source system with a relational database, is considered structured data. With unstructured data (e.g., web data, text data, audio recordings), each instance has its own internal structure (Kelleher & Tierney, 2018), and cannot be easily represented in a table format without transformation.

An illustrative example of structured data from the world of automobiles is shown below in Table 3, where brand, year, customer experience and price, represent the attributes, and each row shows an instance. Attributes come in numeric (e.g., price), nominal (e.g., car brand), and ordinal (e.g., scale of customer satisfaction) forms (Lantz, 2015).

Attributes

Brand	Year	Customer experience	Price in CHF
VW	2012	Very happy	14'800
BMW	2015	Satisfying	30'659
Toyota	2009	Not at all happy	11'356
…	…	…	…

Instances

Table 3 Illustrative example of a dataset and its characteristics based on Lantz (2015)

Corresponding examples for unstructured data would be a collection of photographs of the automobiles, or a set of free-text reviews of each automobile.

For the auditor, journal entries are a fundamental source. They are structured data and consist of numeric and nominal attributes. Depending on the set-up of the audited company, journal entries are created manually or automatically by the system. Journal entries have several attributes, e.g., posting date, currency, monetary amount, debit and credit account numbers (Schreyer, Sattarov, Borth, Dengel, & Reimer, 2018). Each account number is linked to a specific account, which belongs to the chart of accounts. The structure of the chart of accounts and the accounts contained in it are tailored to the specific needs of the particular entity (Debreceny & Gray, 2010). All journal entries are stored in various sub-ledgers, which are linked to the general ledger.

3.2.3 Data quality

Before applying any data analytics tools to a dataset, it is a prerequisite to have the data in place, and to identify its potential shortcomings. Dzuranin and Mălăescu (2016) advise auditors to understand the available data and to get a feeling for its quality and characteristics.

In some cases, data dictionaries are provided, which describe the attributes of the datasets (Han et al., 2012). However, in the absence of such a data dictionary, an understanding of the attributes must be gained in a different way, for example, through interviews and manual analysis. A lack of understanding can sabotage the success of a data analytics project. For instance, it cannot be automatically assumed that even numeric data always has the same semantic meaning (Runkler, 2016). Using various processes to explore and become familiar with the data, including calculating basic statistical measures, can provide the first valuable insights to the analyst (Wirth & Hipp, 2000).

High quality and suitable data analytics can only be achieved using correct and complete data, hence its vital role (Helfert, 2000). If the quality of the data is not assured, the best analytical tools will fail to deliver robust results. This is colloquially known as the "garbage-in/garbage-out" issue. If data quality is insufficient, a short-term solution is to either exclude the incidents with low quality or to carry out manual adjustments. If the errors in the data do not follow any apparent patterns, then beyond a small dataset size, manual adjustments are not a viable option. As Helfert (2000) concludes, high-quality data can only be obtained consistently through the implementation of data quality controls throughout the data warehousing process.

Wang and Strong (1996) developed a widely cited conceptual framework to map all the characteristics defining data quality. This framework is shown in Figure 6. In addition to the factors of data quality, Helfert (2000) also list the recommended measurement methods for checking each quality characteristic group.

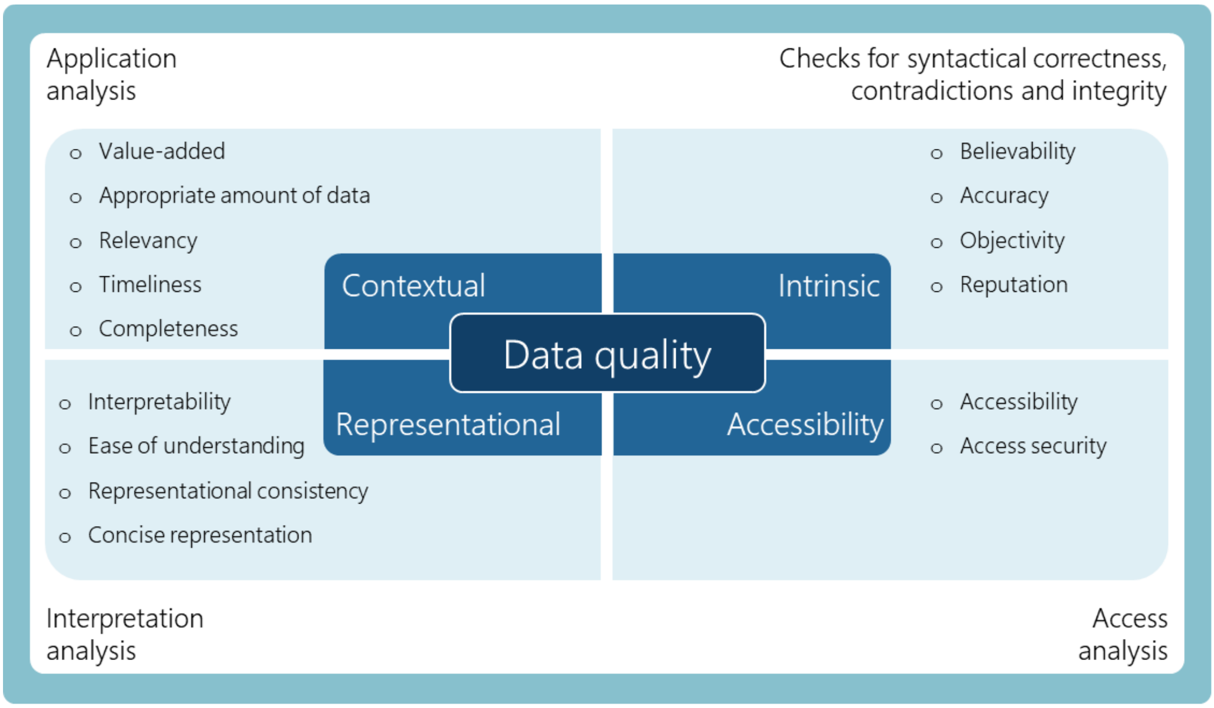

Figure 6 Conceptual framework of data quality based on Wang and Strong (1996, p. 20) and measurement methods based on Helfert (2000, p. 70)

The importance of each characteristic depends on the specific field of application (Mertens et al., 2017). As Han et al. (2012, p. 101) write "data have quality if they satisfy the requirements of the intended use". For auditors, the most crucial data quality attributes are completeness, accuracy, integrity, and reliability (Dzuranin & Mălăescu, 2016) – terms which are found in the ISA standard. Without these attributes, the data is of little use to the auditor, except perhaps as a first red flag.

3.3 Data preprocessing

In the data preprocessing stage, the data is prepared and converted into a form that fits the method chosen for building a model. In some cases, multiple datasets need to be linked together (data integration). Another common preprocessing task is turning unstructured data into structured data (data transformation) (Han et al., 2012).

3.3.1 Missing data entries

Depending on data quality and the intended use of the data, cleaning might be necessary due to noise and inconsistencies (Han et al., 2012). A strategy for dealing with missing attributes of samples needs to be chosen. Various possibilities exist, such as excluding those entries from the dataset, manual replenishment, and inserting measures of central tendency for the attribute. (Han et al., 2012).

The appropriate way of dealing with missing data entries is highly dependent on the distribution of missing data entries in the dataset (Roiger, 2017). While excluding all incomplete data points is often the simplest approach, it might not be a viable option if a specific subset of interest in the dataset is disproportionately affected by omissions.

3.3.2 Data reduction

Kelleher and Tierney (2018) recommend excluding irrelevant or redundant attributes, which could harm the results of the data analysis (data reduction). Additionally, it should be noted that including too many attributes in the analysis increases the probability that methods based on algorithms will uncover spurious patterns, which appear to be statistically significant but are actually meaningless (Han et al., 2012). Several possible methods of attribute subset selection exist (Abbott, 2014; Kotu & Deshpande, 2014; Wu & Coggeshall, 2012). Two basic heuristic methods are briefly explained in the following paragraph.

Forward linear regression selection starts with an empty set of attributes and adds a single attribute from the dataset that yields the highest r^2 (Wu &

Coggeshall, 2012). R^2 is an indicator of the explanatory power of a linear regression that measures how well the attributes explain the values of the dependent variable (Lantz, 2015). In every round, the model goes through all not yet included attributes to find the one that, if added to the model, will give the new model the highest r^2 increase. It does this until adding attributes no longer adds a significant r^2 improvement.

Backward linear regression selection works the other way around. It starts with all attributes included in the model, and in each round, the attribute with the lowest r^2 contribution is eliminated from the set (Wu & Coggeshall, 2012).

A combination of both models is also possible, where the number of attributes used is fixed from the beginning. For each step, the attribute with the highest r^2 is then added, and the one with the lowest r^2 is removed.

Kelleher and Tierney (2018) note how difficult it is to choose the correct attributes for the analysis and advocate an iterative process of trial-and-error experiments. Roiger (2017, p. 52) recommends examining the degree of correlation between numeric attributes because if two attributes "are highly correlated in a positive or a negative direction, only one of the attributes should be selected for data mining. Keeping both attributes is like giving an added weight of importance to a single attribute." However, such corrections must be approached cautiously because the correlation of two attributes does not automatically imply causality. Furthermore, decisions in the preprocessing stage can have a significant downstream impact on the results of the analysis in ways that may not be immediately apparent (Aggarwal, 2015).

Alternatives to attribute subset selection are to reduce the dimensionality of the dataset with a principal component analysis (Han et al., 2012). This method is described in more detail in subchapter 6.2.

3.3.3 Preparing data

In conclusion, many decisions need to be made already during the data preprocessing stage.

When preparing data, before creating a model, model accuracy can be improved not only by reducing attributes but also by combining attributes (attribute construction), multiplication, squaring, and so on (Han et al., 2012).

Factors that influence the dataset size are the minimum required number of instances (rows), the algorithms used, and the number of attributes (columns). Some algorithms, such as neural networks, are much more data-hungry than others, such as linear regressions (Cerny & Proximity, 2001). Finally, a good understanding of the existing data, with in-depth analysis, is crucial. It needs to be understood how the data collection process happened. This helps the data analyst with the issue of dealing with missing values and plausibility checks.

The data preprocessing stage is as critical to the success of the entire process as it is tedious. Roiger (2017) calls it the most-time consuming stage. Although the CRISP-DM sets out a precise sequence of stages, it is common to return to these tasks throughout the entire process. For example, once the first models have been created in the next stage, iterations back to the preprocessing stage are in practice often unavoidable.

3.4 Methods

A carpenter selects the right tool from his toolbox depending on what he wants to achieve, and what kind of material he will be working on. A data scientist picks his method based on the aim of the analysis and the type of data available. The attributes in the dataset help in choosing an appropriate model (Lantz, 2015).

Within this dissertation, five broad categories of methods are discussed, each encompassing several specific approaches:

Data analytics methods	Classification
Rule-based methods	Conventional
Descriptive statistics and visualization	Conventional
Machine learning	Advanced
Natural language processing	Advanced
Process mining	Advanced

Table 4 Data analytics methods for auditors based on Ruhnke (2019)

Machine learning, process mining, and natural language processing are considered advanced methods because they are more challenging to implement (Ruhnke, 2019). Advanced methods are the main focus of this chapter. The conventional methods – rule-based methods, descriptive statistics, and visualization – are briefly reviewed for completeness.

3.4.1 Rule-based methods

For this approach, pre-defined rules are tested on the dataset. The set of rules are created manually by the auditor. The quality of the results greatly depends on the creativity, knowledge, and experience of the person defining the rules.

Rule-based methods are used in auditing for checking hypotheses or identifying suspicious journal entries ("red flags"). Underneath the defined rules lie assumptions about particular risks (Ruhnke, 2019). The objectives of rule-based methods are generally to identify data points in the dataset with fixed rules. Most of the early expert systems (one approach to artificial intelligence) would fall into this category, having explicitly programmed rules defined by human experts (Arnold, Collier, Leech, & Sutton, 2004; Eining, Jones, & Loebbecke, 1997).

Typical examples for rule-based methods are approaches designed to detect journal entries entered on the weekend or during the night in the dataset. Other examples are to check for negative inventory numbers or duplicate journal entry numbers. Slicing and dicing data with filters and sorting, as is commonly done in spreadsheet software, can also be considered a part of rule-based methods.

The rules can come from known business relations or statistical facts. For example, Benford (1938) observed that real-life sets of numerical data composed of four or more digits follow a specific distribution, in which the probability of observing one as the first digit is almost six times higher than observing a seven. By identifying this distribution as generally applicable, he formulated "Benford's law", named after him. While it has its limitations and cannot be applied in every area (e.g., heights of adults) (Cleary & Thibodeau, 2005; Druică, Oancea, & Vâlsan, 2018), it has been shown in several scientific publications to be useful in the external auditing context (Nigrini, 1999; Nigrini & Mittermaier, 1997). Suh and Headrick (2010) recommend having a minimum sample size of 1'000 observations to conduct a first- and second-digit analysis with Benford's law. With fewer observations, only a single-digit analysis can be performed.

Rule-based methods are straightforward. This is a strength because the results are easy to interpret, and the auditor can understand the approach. However, the quality of the output relies heavily on the skill of the person defining the rules. Because the rules often include explicitly set thresholds, it can create blindness to insights available outside of the defined rules (Jans et al., 2011;

Westermann & Spindler, 2017). Furthermore, as the set of defined rules grows, it quickly becomes challenging to maintain them.

3.4.2 Visualization and descriptive statistics

Han et al. (2012, p. 82) define the aim of data visualization as "to communicate data clearly and effectively through graphical representation". Today, most graphical representation is done with computer support. In practice, the most commonly used tool is Excel. However, this tool has its limitations, and software also exists for more complex visualizations such as PowerBI or Tableau. According to Liu (2014), basic data visualization tools are capable of generating diagrams such as bar charts, box plots, histograms, and scatter plots. More advanced tools can produce visualizations such as heat maps or provide dashboards. The latter typically include graphical, text-based and tabular information all combined and displayed on one screen (Abts & Mülder, 2017).

An example of an audit dashboard is provided in Figure 7.

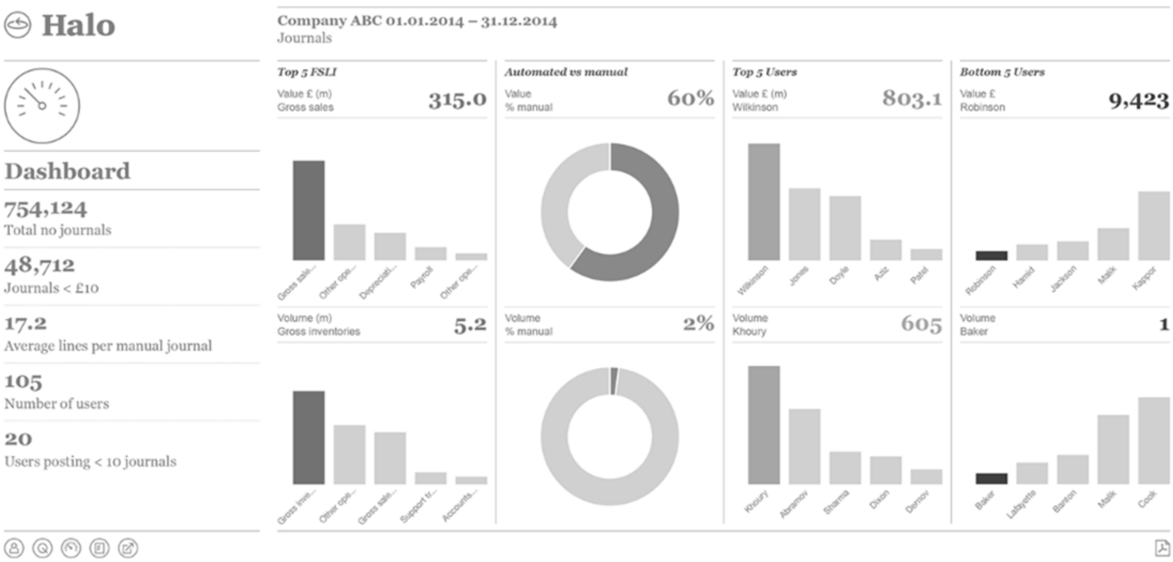

Figure 7 Illustrative example of audit dashboard for journal entries retrieved from PwC (2012)

After uploading the journal entries from the client, various facts are displayed on the dashboard, such as the number of journal entries, top and bottom users, and automated versus manual journal entries.

According to Mann (2001, p. 3), "descriptive statistics consists of methods for organizing, displaying, and describing data by using tables, graphs, and summary measures". This includes frequency distributions, calculating lower and upper limits, looking at the data distribution, and measures of central tendency (median, mean, mode), to name a few. Some descriptive statistics techniques will have typically already been employed during the data understanding stage.

Visualization and descriptive statistics are so widely used today, that we take them for granted, to the degree that it even seems strange to mention them as explicit data analytics methods. However, it is critical to realize that the ability to quickly gain summary statistics without manual calculation work is the basis for many of the more advanced methods. In the same sense, data visualization is a recurring theme throughout. Vision can provide intuitive access to data in a way that large arrays of numbers themselves cannot (Conway & White, 2012). Therefore, effective visualization is incredibly useful, not only for finding patterns in data and extracting meanings from it, but also for communicating these insights and findings.

3.4.3 Machine learning

Machine learning can be explained as algorithms, which learn to identify and extract patterns from the presented data (Kelleher & Tierney, 2018). It is a subcategory of artificial intelligence. Some machine learning algorithms (e.g., logistic regression) use traditional statistical analysis but focus on finding mathematical equations to predict an outcome or a classification (Dickey, Blanke, & Seaton, 2019).

Machine learning has produced several success stories over the past few years. For example, machine learning algorithms are used in e-mail spam filters, for identifying fraudulent credit card transactions, customer segmentation in targeted advertising campaigns, and in medicine for tumor detection (Lantz, 2015).

As Kelleher and Tierney (2018, p. 101) write, "The real challenge in using ML [machine learning] is to find the algorithm whose learning bias is the best match for a particular data set", because each algorithm generalizes differently from a dataset. This phenomenon is also known as learning bias, selection bias or modeling bias.

When performing data analytics with machine learning algorithms, several models are usually built in this stage, and the performance, as well as the usability, are compared afterwards (Provost & Fawcett, 2013). A golden rule is to split the available dataset into training, validation, and testing sets. Commonly used ratios are: 50:20:30 or 40:20:40 (Kelleher & Tierney, 2018).

Based on Kelleher and Tierney (2018), the datasets are explained as follows. The training set is used to train the models. Then the validation set is employed to compare the performance of several models on new data not used for training, to find the best one. Once the most appropriate model has been selected, the validation and training sets are combined to train the model on a larger dataset. The testing set is then used to evaluate the performance of the final model.

It can be difficult to collect enough data even for the training set. Still, it is crucial to reserve a portion of the overall data for exclusive use in validation and testing purposes (Provost & Fawcett, 2013). If the same data is used for training that would later be used for validation and testing, there is a high risk of obtaining a model that is overfitted to the training set. Such a model is not generalizable to instances outside its original training set and is therefore useless for making predictions.

Lantz (2015) divides the universe of machine learning algorithms into two categories: supervised learning and unsupervised learning[6].

For supervised methods, the dataset includes "labeled" examples with target information (e.g., fraud, nonfraud, bankrupt, nonbankrupt). The target information is typically the question of interest. Each labeled example comes with a set of attributes (e.g., financial ratios) that are used to predict the target information. In other words, supervised learning is based on learning from examples and then applying what has been learned to a larger dataset.

[6] A combination of both methods exists (semi-supervised learning), which is not discussed in this dissertation.

For unsupervised learning methods, the datasets used are without labeled target information. Unsupervised learning algorithms do not know the purpose of learning and try to discover patterns on their own (Provost & Fawcett, 2013). In other words, unsupervised learning is about detecting structure and relationships in a dataset with a less predetermined focus. Unsupervised learning methods include clustering, association rules, and unsupervised neural networks.

Figure 8 outlines a number of supervised and unsupervised learning methods, which will be described in detail in the subsequent sections. Only those methods which are mentioned in the empirical sections of this work (chapters 4 to 7) are presented, representing a curated selection from the larger collection of machine learning methods in existence.

Supervised learning
Nearest neighbor
Naïve Bayes and Bayesian belief networks
Decision trees and random forests
Multiple instance learning
Linear and logistic regression
Artificial neural networks / deep learning
Support vector machines
Genetic algorithm
Ant colony optimization
Unsupervised learning
Association rules
Clustering
Unsupervised neural networks

Figure 8 Covered machine learning methods in this dissertation based on Lantz (2015)

3.4.3.1 *Supervised learning*

Supervised learning includes classification, regressions, and black-box methods such as neural networks and support vector machines. Instead of using just one machine learning method, combinations of methods can also be helpful. This is called an ensemble, or voting ensemble, as each algorithm within the ensemble votes about the classification of each instance (Knox, 2018). Based on the majority vote, the instance is assigned to the corresponding class. Using the ensemble method usually improves model performance but comes at the cost of increased complexity.

Three examples of ensemble-based methods are bagging, boosting, and stacking (Larose & Larose, 2015). Boosting and stacking can improve the accuracy rates of models, while bagging reduces the variance of the ensemble classifiers, which is useful when using unstable classification models (e.g., decision trees) (Knox, 2018; Larose & Larose, 2015).

A final caveat concerns the need for labeled data that is a prerequisite for the application of any supervised learning method. If a method needs a large dataset to achieve reliable results, then the effort of labeling the raw data can be time-consuming and is a substantial cost-driver of the entire data analytics project.

3.4.3.1.1 Nearest neighbor

The nearest neighbor methods are called "lazy learners" because they classify given examples based on the closest data points (neighbors) of the known examples from the training set. To measure the distance between data points, the Euclidean distance is typically used, but alternatives are available (Han et al., 2012). Nearest neighbor is performed by applying an algorithm such as k-nearest neighbors (k-nn), which has been known since the 1950s (Han et al., 2012). A key input is the choice of k, which is set by the creator of the model and indicates how many "neighbors" (data points) are included in the classification of the new examples. An example illustrated in Figure 9 indicates the effect that the choice of k can have on the output of the model.

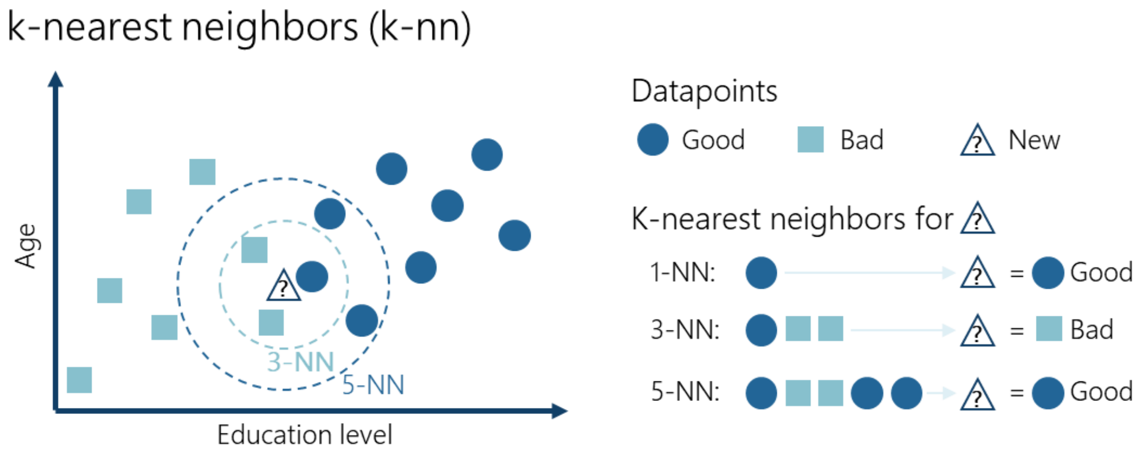

Figure 9 Example of k-nn based on Moreira, De Carvalho, and Horváth (2019)

In practice, finding the best value for k is an iterative process and is achieved by running the model with different k-values and comparing the results. Some research tackles this issue by creating algorithms that determine the ideal value of k based on specific criteria (Shafeeq & Hareesha, 2012). To make sure the algorithm does not include a bias towards the various distance scales given, Lantz (2015) suggests rescaling (e.g., normalizing) all variables to make sure each attribute contributes proportionally the same to the distance calculation. For this type of classification, nominal features are either not included in the model or need to be transformed to a numerical value.

3.4.3.1.2 Naïve Bayes and Bayesian belief networks

The Naïve Bayes algorithm has its roots in statistics and probability theory (Kotu & Deshpande, 2014). The method's classification output is an estimated likelihood. Its name is derived from its characteristic of making naïve assumptions about the independence between variables, which may not always be accurate and is an inherent limitation of the model.

As the model calculates the probability of variables, it can be that the dataset contains no examples for some combinations. In cases where a combination has not occurred in the given dataset, it is recommended to apply the Laplace estimator, which adds a small number to ensure all features have a nonzero probability of occurring (Lantz, 2015).

A related but more advanced method is Bayesian belief networks, which does not assume conditional independence of the attributes (Han et al., 2012). These networks allow for the definition of the conditional independencies between subsets of attributes with parent and descendant relations. Those relations are either defined or inferred from the data by the creator of the model. According to Han et al. (2012), defining the independencies of the attributes can significantly improve the learning rate. The model can return either a single class allocation or a probability indication for how likely it is that the example belongs to a particular class.

3.4.3.1.3 Decision Trees and Random Forests

Decision trees find classification rules that divide the training dataset based on common features for the target attribute. Numerous algorithms, such as the C5.0, exist to build decision trees (Lantz, 2015). The algorithms are optimization functions that iterate over the dataset to find the set of rules that result in the least number of misclassifications.

Decision trees are related to and similar to rule-based methods. The essential difference is that in decision trees, which is based on machine learning, the rules are learned from the data, whereas in rule-based methods, the rules are manually defined by human experts.

Figure 10 shows an example of a decision tree for credit card approvals. Decision trees are straightforward and easy to understand, with a visualization of the rules and the simple if-then-else logic. Additionally, decision trees can handle combined numeric and nominal data input (Kelleher & Tierney, 2018).

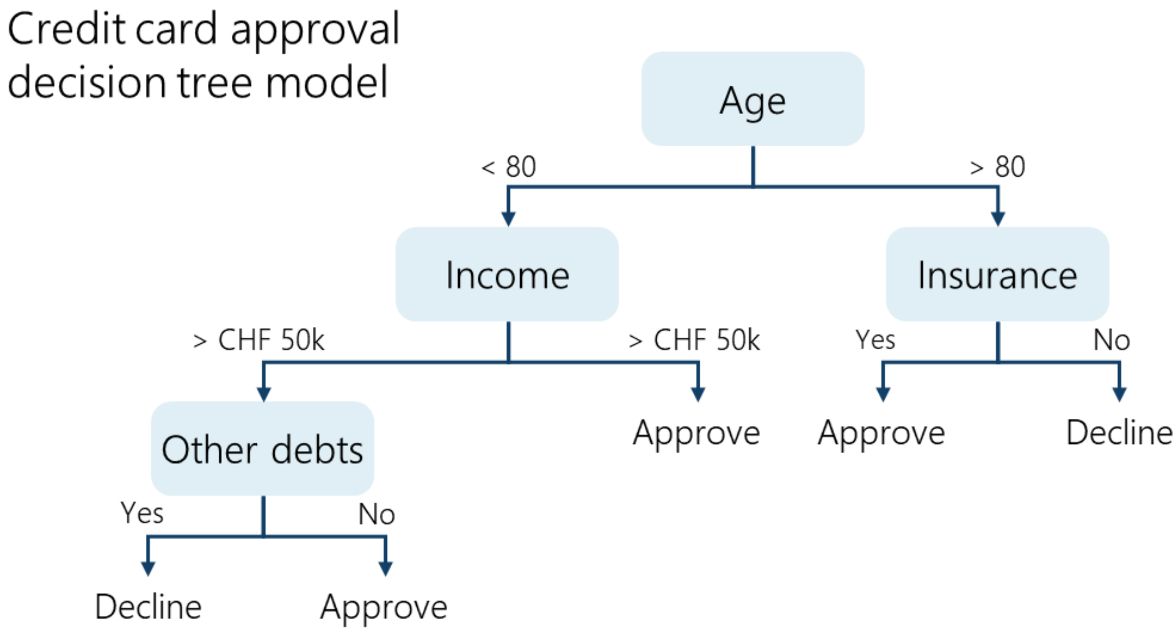

Figure 10 Example of a decision tree based on Roiger (2017)

An ensemble method called “random forests” focuses on improving the performance of decision trees. The naming of the method is derived from the fact that combining various (decision) trees yields a forest. For classifying an instance, a vote is collected from each tree, and the most popular class allocation is returned (Wu & Coggeshall, 2012). The trained decision tree model can be effectively applied to a dataset. As Figure 10 shows, the rules can be easily derived from the decision tree, and the application of such rules is computationally efficient even when applied to large datasets. The results of classification models are expressed in terms of probability or likelihood scores indicating how likely something is to happen, based on the information provided.

3.4.3.1.4 Multiple instance learning

When using multiple instance learning, the instances are put into labeled “bags” (Babenko, 2008). The bag labels are restricted to binary options (negative/positive), for example, fraud and nonfraud cases (Kotsiantis & Kanellopoulos, 2008). The algorithm looks at the similarities of the instances and the label of the bag. The algorithm creates hypotheses based on the accumulated knowledge to be able to classify unseen bags or instances (Babenko, 2008).

3.4.3.1.5 Linear and logistic regression

Regression models try to create a function that explains and predicts the target variable (dependent variable) with the given attributes (independent variables) (Kotu & Deshpande, 2014). Regression models search for the best parameters for the independent variables to reduce the amount of unexplained variance in the dependent variable (measured by r^2) to a minimum. The purest form of regression is linear regression, which works only with numeric data.

An equation describes the relationship between the independent variables and the target variable. The equation reveals which variables have the strongest influence on the target/dependent variable (Runkler, 2016). While the linear regression model is limited to modeling a straight line, logistic regression has – due to the logistic function – the ability to model more complex dependencies.

3.4.3.1.6 Artificial neural networks / deep learning

Artificial neural networks are inspired by the learning process of the human brain. The synapses between neurons get stronger when external stimuli are introduced (Aggarwal, 2015). The neurons in neural networks, called nodes, are connected. Those nodes are arranged in groups known as layers. Many different types of network topologies exist. The simplest one is a single layer network, but this can be increased to multiple hidden layers or multiple output nodes for more complex models. To illustrate how neural networks function, a multiple-layer network is briefly described (Lantz, 2015) and shown in Figure 11.

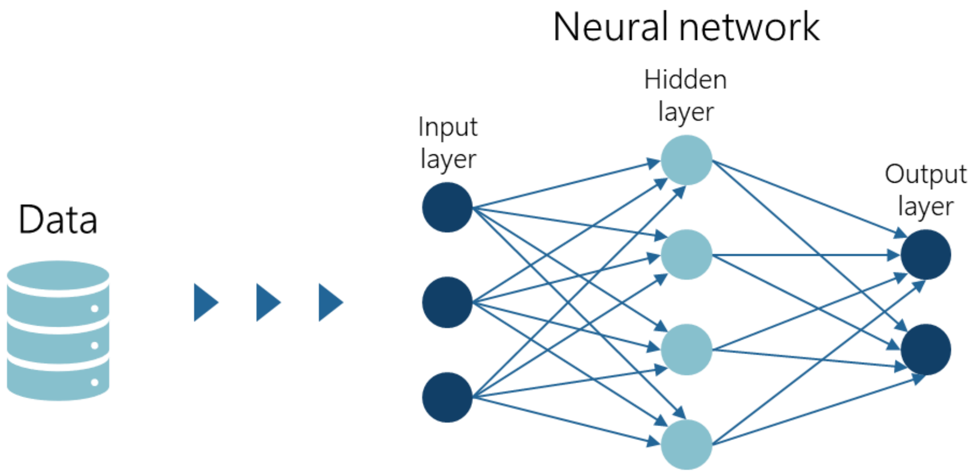

Figure 11 Example of multi-layer neural network based on Aggarwal (2015)

The input nodes, also called the input layer, receive the dataset, and pass it on to the connected computational nodes, named hidden layers. The final nodes in the model are called the output layer. The number of input nodes is defined by the number of attributes in the dataset, and depending on the chosen network algorithm, the model creator sets the number of hidden layers. The number of output nodes depends on the number of results.

Each connection between the nodes ("synapses") has a different strength, which is indicated by weights. As the neural network contains no knowledge at the beginning, the weights are set at random values (Lantz, 2015). The calibration of those weights depends on the architectural set-up. Either forward- or backward-propagation can be chosen, though the latter is the more popular, according to Abbasi, Albrecht, Vance, and Hansen (2012). With forward propagation, the weights are adjusted with the flow of data through the system (Runkler, 2016). In backpropagation, the suggested output is compared with the actual value. In cases of difference, a signal is sent back to the neurons which leads to the weights being adjusted. During the learning process of the neural network, those weights are repeatedly readjusted to minimize incorrect predictions.

As Aggarwal (2015, p. 326) writes, "the key to the effectiveness of the neural network is the *architecture* used to arrange the connections among nodes", which indicates that the chosen network topology is a crucial part of the model. It is possible to include "feedback for learning" in the architectural structure, ensuring the network continues to learn (Kantardzic, 2011).

A deep neural network, which carries out what is commonly known as deep learning, is a type of neural network with multiple hidden layers (Lantz, 2015). This method is recognized as being extremely "data-hungry", implying that a large amount of data is required for training (Issa et al., 2016). However, no definite minimum quantity can be found in the literature, as this depends on the input data itself.

3.4.3.1.7 Support vector machines

Support vector machines (SVM) use a unique approach to classify linear and non-linear data (Roiger, 2017) by transforming the data into a multidimensional surface (Aggarwal, 2015). Within this higher dimension, SVMs can separate the data into classes (Han et al., 2012). The SVM essentially searches for an optimal dividing nonlinear or linear line between the data points that ideally creates a neat border between the given classes in each dimension (Aggarwal, 2015). SVMs are an example of "eager learners" because they learn a classification from the first dataset and hence are ready and "eager" to classify unseen data (Han et al., 2012).

3.4.3.1.8 Genetic algorithms

Genetic algorithms are inspired by Darwin's "survival of the fittest" biological evolution theory (Aggarwal, 2015). During the learning stage, only the "fitness relevant" features are selected to contribute to the output, which is an advantage for datasets with many features (Hoogs, Kiehl, Lacomb, & Senturk, 2007). The evaluation function for determining fitness is set by the model creator (Roiger, 2017).

A significant advantage of genetic algorithms is the "capability to learn class boundaries that are non-linear functions of multiple variables, allowing for solutions that cannot be achieved by linear methods." (Hoogs et al., 2007, p. 42). Based on the chosen application field, genetic algorithms classify or cluster their output. Genetic algorithms can also be used in conjunction with other algorithms such as neural networks (Anandarajan, Picheng, & Anandarajan, 2001) or during the data preprocessing stage (Roiger, 2017).

3.4.3.1.9 Ant colony optimization

Another bio-mimicry algorithm is ant colony optimization, which takes inspiration from a simple algorithm first observed in nature among ants. Ant colony optimization is a swarm intelligence method which mimics the phenomenon of ants leaving pheromones on their way to and from collecting food (Katsis, Goletsis, Boufounou, Stylios, & Koumanakos, 2012), thus making a trail for other ants to follow. The more ants that choose the same way, the stronger the pheromone trail will be. Due to the self-reinforcing effects of this feedback cycle, ant colony optimization eventually arrives at an optimized path.

The principles underlying this algorithm can be generalized and applied to a broad range of optimization problems. The model creator decides how many virtual ants will be employed. From the results of the algorithm, simple "if-then rules" can be extracted.

3.4.3.2 *Unsupervised learning*

In contrast to supervised learning, datasets for unsupervised learning do not need labeled data. Based on the characteristics in a dataset, a computer develops clusters and assigns each instance to the corresponding cluster. However, coming to an opinion about what each cluster represents is a judgment left to the model creator.

3.4.3.2.1 Association rules

"Association analysis measures the strength of co-occurrence between one item and another" (Kotu & Deshpande, 2014, p. 195), meaning that the algorithm tries to find rules or patterns within a dataset. This method is also known under the name "market basket analysis", as it became famous for analyzing supermarket data (Lantz, 2015). In a nutshell, it yields results of the pattern: "customers who buy x also buy y". Several algorithms exist to perform association analysis (Aggarwal, 2015), such as the apriori algorithm, which is used in subchapter 6.4.

3.4.3.2.2 Clustering

Clustering automatically classifies a dataset into a defined number of groups. Clusters are formed based on a similarity measure, meaning data points with similar attribute characteristics belong to the same group. Similarity measures can be of a qualitative or quantitative nature. Once the similarity measures are defined, the distances between data points are measured. The calculation depends on the scale type of the attributes (Moreira et al., 2019), with the most similar ones having the smallest distances between them and vice versa. In some instances, it may be helpful to normalize the attributes (replacing absolute values with their distance from the mean as measured in standard deviations) before implementing clustering so that differences in scaling do not distort the results.

Various clustering algorithms exist, of which the most well-known is k-means clustering (Kantardzic, 2011). K-means is a centroid-based clustering technique. Each cluster has a centroid (point in the middle of the cluster), and each new instance is assigned to the closest cluster, measured by the distance to the next centroid (Han et al., 2012). K represents the number of clusters in the given dataset and is an input determined by the model creator. An example of k-means clustering is shown in Figure 12.

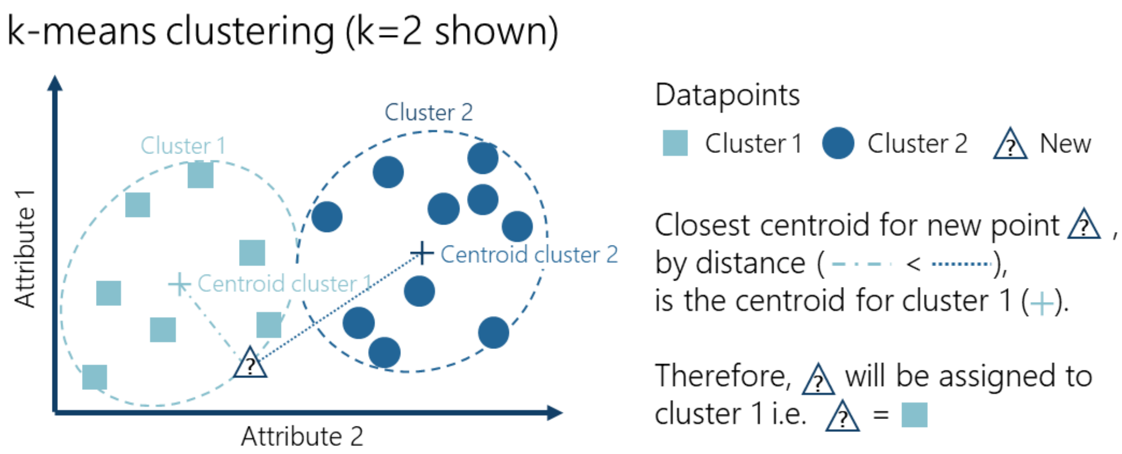

Figure 12 Example of clustering based on Moreira et al. (2019)

While k-means is easy to understand and is a suitable clustering method to start with, more sophisticated clustering algorithms are available today (Lantz, 2015). One such method is fuzzy clustering. According to Runkler (2016, p. 109), this method "works well if the clusters are well separated and do not contain inliers or outliers". Another alternative is the latent class clustering algorithm, which allows for overlapping clusters (Jans, Lybaert, & Vanhoof, 2010).

3.4.3.2.3 *Unsupervised neural networks*

While neural networks have been discussed before under supervised learning methods, there are also variations of this method that can work with unlabeled input data. Self-organizing maps are an example. According to Kiang (2001, p. 161f), self-organizing maps are "a special type of neural network that can learn from complex, multi-dimensional data and transform them into visually decipherable clusters". Another example is autoencoder neural networks, which will be covered in detail in subchapter 6.3.

3.4.4 Natural language processing

Natural language processing (NLP) represents the computational ability to process human natural language, which helps extract insights from text (Thanaki, 2017). For processing natural language, as opposed to programming languages, various methods, including machine learning, are used in combination.

Natural language processing contains several preprocessing steps, which improve the ability to parse the structure and intricate patterns inherent in human languages. Depending on the intended area of use for NLP, this includes various language skills such as semantic, lexical, syntactic, and morphological analysis (Thanaki, 2017). For conducting these analysis, there is the option to either use available open-source tools, buy proprietary tools or develop customized tools in-house (Hardeniya, Perkins, Chopra, Joshi, & Mathur, 2016).

The input data consists of text, which needs to be processed first and turned from unstructured data into at least semi-structured data with feature engineering. During this process, features are generated or extracted from the given text (Thanaki, 2017). This step can include activities such as breaking raw input text into words (tokenization), reducing verbs to common roots by removing tenses (stemming), and eliminating punctuation. After breaking up a text into words (tokens), each token receives a part-of-speech label (Ghosh & Gunning, 2019). For example, "today is busy", is split up into the tokens "today", "is", and "busy". Each token is assigned a part of speech label: "today"

- noun; “is” - verb; and “busy” - adjective. However, interpreting the content in terms of meaning can be difficult. For example, “today” can have several meanings such as on this day or nowadays. This can lead to ambiguity, which makes it necessary to carry out a word sense disambiguation process to make sure the correct meaning is assigned to the token. This basic example shows how challenging the preprocessing phase for natural language processing can be. Even though tasks in the processing stage can be automated, the selection and careful application of preprocessing steps for natural language processing require an experienced data scientist.

Once the feature engineering is completed, the preprocessed textual data is fed to a machine learning algorithm and applied to create models to answer the business question at hand. Successful examples of NLP have already been incorporated into our daily lives in the form of spelling correction in text editing software, search engines, and speech engines (Hardeniya et al., 2016).

3.4.5 Process mining

Van der Aalst and Weijters (2005, p. 235) define the basic idea of process mining as “to extract knowledge from event logs recorded by an information system” (e.g., ERP-system). Usually, the results of process mining are visualized as a process flow diagram (see Figure 13 on the next page). An event log in this context is “a chronological record of computer system activities” (Jans, Alles, & Vasarhelyi, 2013, p. 4). It follows that this method is only viable if event logs are given.

Compared to other data analytics methods, process mining is a young research discipline (Van der Aalst, 2011). It is a mix of machine learning, process modeling, and process analysis. In the context of auditing, process mining facilitates a comparison between the design of a process versus its actual operation, and enables the auditor to visualize the actual process flow (Jans et al., 2013).

According to Jans, Alles, and Vasarhelyi (2014), the following four attributes must be extracted for each event in order to perform process mining:

1. Activity taking place during the event (e.g., sign, create purchase order)
2. Process instance of the event (e.g., invoice, goods receipt)
3. Originator of the event (user or system automatically performing an activity)
4. Timestamp of event occurrence (automatically performed by the system)

Process mining algorithms are based on process models that are abstractions of the actual processes being analyzed (Werner & Gehrke, 2015). Various process model notations can be used to make the processes understandable.

While process flows within companies are mostly explicitly defined, some flexibility is required in practice to allow smooth operation of the business (Jans et al., 2014). Such flexibility comes at the cost that diverse and less-structured process executions exist. This can result in less structured processes, called "Spaghetti" processes (Günther & Van der Aalst, 2007), and is shown in Figure 13 on the next page on the left-hand side. Van der Aalst (2011) suggests as a rule of thumb an 80% fitness threshold for the differentiation between structured ("Lasagne") and unstructured ("Spaghetti") processes. This implies that more than 80% of the processes should be carried out as planned. Günther and Van der Aalst (2007) developed a Fuzzy Miner algorithm to deal with the complexity of "Spaghetti" processes. Additionally, the authors implemented an edge cutoff parameter, which regulates the mapping of the processes. When the parameter is set to zero, all processes are shown (see Figure 13 left side), but when increased, the visualization displays fewer details by removing infrequent process occurrences from the visualization.

In Figure 13, on the right side, an edge cutoff of 0.4 was used. Such a feature can help users to bring more structure into a chaotic process model.

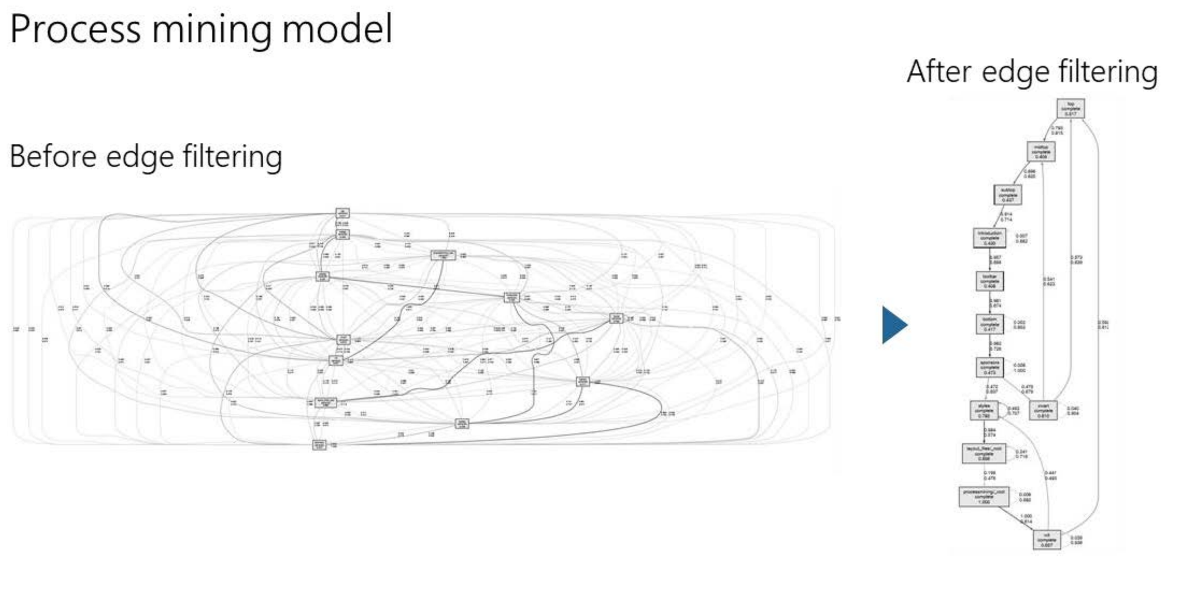

Figure 13 Example of process model retrieved from Günther and Van der Aalst (2007, p. 339)

Process mining is especially handy for auditors to create flow-charts and understand how the actual execution of the internal controls proceeded during the year under audit. As Jans et al. (2014) wrote:

> *Although processes are mostly prescribed to take place according to a designed process model, some room for flexibility is necessarily built into the business's ERP-system. This flexibility is designed to allow for the numerous deviations from the designed process model that are often required in practice for smooth operation of the business, constraints that would otherwise interrupt the process flow too frequently (p. 1755).*

As an example of this, the authors name the three-way match, which is the reconciliation of purchase order, invoice, and goods receipt. In case of deviations, such as between purchase orders and goods receipts (e.g., due to transportation costs not included in the purchase order), goods would be rejected, which can create delays in the production process. Tolerance levels are implemented in the ERP-system to deal with such issues.

3.5 Evaluation

Once a method has been selected and a model is calculated, the results must be interpreted, and the quality of the results needs to be evaluated. "The purpose of the evaluation stage is to assess the […] results rigorously and to gain confidence that they are valid and reliable before moving on" (Provost & Fawcett, 2013, p. 31).

It is essential to define adequate evaluation criteria already in the business understanding stage, to ensure the models created meet the deployment requirements. Although there are several mathematical evaluation methods, especially for the more advanced methods, all of them have limitations. None of them are a substitute for human judgment and understanding when it comes to interpreting and evaluating the results of a model or data analytics approach.

In cases where the model performance is not satisfying, various methods exist to improve it (e.g., standardizing or normalizing the data).

3.5.1 Evaluation of conventional methods

For conventional methods – rule-based methods, visualization and descriptive statistics – evaluation generally lies in the hands of the creator, especially concerning the accuracy of the results. However, for the advanced data analytics methods, more sophisticated evaluation criteria are relevant as well.

3.5.2 Evaluation of machine learning methods

The prerequisite for evaluating the results of machine learning methods is the careful management of the dataset. As mentioned, the golden rule for the evaluation of the model performance for machine learning methods is to split the data (Kelleher & Tierney, 2018).

By splitting a dataset appropriately, the generalizability of the model can be enhanced, and at the same time, the dangers of overfitting and underfitting are reduced. Overfitting describes the tendency of the model to memorize the data seen in the training set without uncovering underlying relationships (Provost & Fawcett, 2013). Underfitting is the opposite effect, where the model does not include essential features and consequently lacks the precision required for being useful.

Depending on the machine learning model applied, various criteria exist to evaluate the performance. Especially for supervised machine learning, more comprehensive evaluation criteria are possible as the results can be compared with existing target values. Roiger (2017) recommends applying the evaluation criteria to the results of the validation testing set.

Table 5 displays an illustrative list of evaluation criteria. The precise criterion depends on the specific type of machine learning method employed.

Evaluation criteria	Application	Definition
Area under the curve (AUC)	Classification	Range: 0.5 – 1, where 0.5 means no predictive ability and 1 resulting in a perfect prediction every time
Accuracy	Classification	Range: 0 – 1, generally expressed as percentage of correct classification outcomes
Confusion matrix	Classification	Matrix showing the incorrect and correct prediction of the classes (see also Table 6)
Error rate	Classification	Percentage indicating the mis-classification rate
Sensitivity	Classification	Percentage showing the true positive rate
Specificity	Classification	Percentage showing the true negative rate
(Adjusted) R-squared	Regression Models	Range: 0 – 1 , where 0 means the model is useless and 1 indicates a perfect predictor
Confidence	Association rules	Degree of certainty of the detected pattern
Support	Association rules	Percentage of data to which the pattern applies

Table 5 Illustrative list of evaluation criteria for machine learning models based on Han et al. (2012) and Roiger (2017)

For association rules, confidence and support are not evaluation criteria per se, but are used instead as thresholds to prevent an overload with unusable patterns.

Table 6 on the next page shows an example of a confusion matrix based on fraud.

<table>
<tr><td colspan="2"></td><th colspan="2">Predicted class</th></tr>
<tr><td colspan="2"></td><th>Positive</th><th>Negative</th></tr>
<tr><th rowspan="2">Actual class</th><th>True</th><td>Fraud firm is classified as a fraud firm</td><td>Nonfraud firm is classified as a nonfraud firm</td></tr>
<tr><th>False</th><td>Nonfraud firm is classified as a fraud firm (Type I)</td><td>Fraud firm is classified as a nonfraud firm (Type II)</td></tr>
</table>

Table 6 Example of a confusion matrix for supervised machine learning methods applied to fraud detection based on Perols (2011)

Depending on the defined use case for data analytics, misclassification costs might play an essential role and may need to be integrated into the model from the beginning (Dbouk & Zaarour, 2017). One solution is to enrich the confusion matrix with the costs of misclassification.

When the model achieves an accuracy rate of around 50%, a user achieves the same result by just randomly guessing or throwing a coin. Such a low accuracy rate could also be a sign of "underfitting", meaning that the model did not pick up on patterns in the data (Dickey et al., 2019). However, exceptionally high accuracy rates do not guarantee a well-functioning model. It could be a sign that the model picked up on noise in the training data, which is known as "overfitting". To prevent this phenomenon, a model should never be trained and tested on the same data.

Another extreme case is when almost all samples can be classified into the same class, and a high accuracy rate is achieved. For example, if two classes exist (e.g., fraud and nonfraud) and the dataset consists of 99% nonfraud examples and 1% fraud examples, the machine learning method can simply assign all instances to nonfraud class and be correct 99% of the time (Finlay, 2017). In such a situation, it is important to question if it is worth building a model at all (base rate fallacy). A proactive approach to this problem is essential during the data preprocessing and modeling stage. There might be trade-offs between the accuracy of the results and their interpretability.

Unsupervised machine learning methods have the advantage that the data does not need labels. However, this advantage comes at the cost of having no objective indicators to measure their performance. A qualitative measure of the worthwhileness of this trade-off might be the usefulness of the results or the gut feeling of the modeler (Lantz, 2015).

Roiger (2017) points out that supervised and unsupervised methods complement each other when it comes to the evaluation of models. For example, the output of a clustering technique can be used to provide a label for the data. In the next step the labels can be tested with a classification algorithm such as k-means.

3.5.3 Evaluation of natural language processing

For natural language processing, the results are evaluated based on the outcome produced (Hardeniya et al., 2016). The outcome can be compared with predefined metrics or checked manually. Natural language processing tries to mimic human language use so the outcome can be benchmarked against the results of a human being. As natural language processing makes use of machine learning algorithms, the evaluation technique corresponding to the algorithm used can be applied to evaluate the model performance. For example, when using supervised learning methods, evaluation metrics, such as the accuracy range or a confusion matrix, can be applied (Ghosh & Gunning, 2019).

For natural language processing, there are also evaluation tools that evaluate the preprocessed data, such as part of speech taggers, stemmers, and morphological analyzers (Hardeniya et al., 2016).

3.5.4 Evaluation of process mining

Van der Aalst (2016) defines four quality dimensions for process mining: fitness, simplicity, generalization, and precision. According to the author, the four criteria should be balanced. Perfect fitness of a model is reached if all traces of the event log are displayed in the process model. The simpler a model is, the easier it is to handle for the user. A good process model is also generalizable but still precise. This mirrors the overfitting (generalizable) and underfitting (precise) criteria, as applied to machine learning algorithms (Provost & Fawcett, 2013).

3.6 Deployment

Deployment is the final stage of the CRISP-DM framework and focuses on the implementation of the designed data analytics methods in a model or tool that can be used. During the business understanding stage, deployment criteria were collected to ensure the developed model fits the intended area and fulfills the pre-defined requirements. Now the created method is put into action to realize a positive return on investment (Provost & Fawcett, 2013).

The success of the deployment depends on many factors. Along with the seamless technical integration of the developed data analytics model, the training of the personnel using it is essential. This includes both initial training and the necessary ongoing support for any arising issues (Chamoni & Gluchowski, 2017).

Once the model has been implemented, regular check-ups and small scale adjustments will be needed to ensure ongoing functionality of the model (Chamoni & Gluchowski, 2017). “Even the most accurate and effective models don’t stay effective indefinitely. Changes in behavior due to new trends, incentives, or disincentives should be expected” (Abbott, 2014, p. 367). Such changes trigger the maintenance of the deployed methods and restart the whole data analytics process.

3.7 Limitations of data analytics

The previous subchapter illustrates how data analytics is carried out and what can be achieved. This subchapter deals with the limitations of data analytics, especially in an audit context.

Data analytics processes only digitally available data. However, there is still a lot of data in companies that is not electronically available or not in a format that makes it practical to include in a data analytics project. This can be quite a burden for a data analytics project, as essential data may be not included in creating the model or will require additional preparation effort. This is also a critical point for auditors when performing internal control tests. Manually performed checks will not be captured by process mining and will, therefore, need to be considered separately during an internal control audit.

A common misconception is that models can be trained with just about any data to produce usable results (Provost & Fawcett, 2013). The right selection of data is often overlooked, and the colloquial phrase "garbage in/garbage out" applies.

Using more advanced data analytics models has become easier with more user-friendly software (Kelleher & Tierney, 2018). This comes with the risk that people without knowledge of the theoretical underpinnings make mistakes when creating advanced data analytics models. They may not even realize when they have made these mistakes, since every model or algorithm has inherent limitations that are frequently not obvious. Add to this the fact that decisions made during the data preprocessing stage can have implications which even experienced practitioners occasionally fail to anticipate correctly.

When building a model, it must be kept in mind that the data itself is a product that involved human decisions or interactions and has inherent limitations such as incompleteness, errors, noise, and data uncertainty (Han et al., 2012; Provost & Fawcett, 2013). Noise and errors, especially, can lead to a model making wrong predictions. Besides, it is not certain that a data query always includes the same type of data. The data itself can be malleable.

Despite these limitations, there is now a large data analytics toolbox available that offers many new opportunities for auditors. In the following chapters, there will be an analysis of which of these tools are being used or are usable for the practice of external auditing.

4 Applications identified in research

4.1 Research methodology

The research methodology aims to answer research question 1:

Which applications of data analytics methods have been identified for external auditing in research?

The chosen approach is to carry out a literature review to get an overview of recent developments and "a complete understanding of the relevant topics" (Hair, 2011, p. 90). Literature reviews are a suitable instrument for identifying new forms of applications (Eriksson & Kovalainen, 2016). The literature review also provides a basis for the next three research questions. The approach for carrying out the literature review is a combination of the methods of Hair (2011) and Knopf (2006). This yields the following collection process:

1. Define sources of information
2. Define keywords and inclusion criterion
3. Start collecting the literature
4. Review the identified sources and apply the inclusion criterion

4.1.1 Data source and data collection

As recommended by Knopf (2006), a broad spectrum of sources was considered for the data collection. By doing so, the results of the literature review encompass the latest developments in the research area going beyond practice or academic sources only. Therefore, books, journal articles, works of standard setters, dissertations, and newspaper articles are included.

The literature review includes sources published from 2000 onwards. This corresponds to the timeframe within which data analytics methods started to become viable (Runkler, 2016). Additionally, what was state-of-the-art before the turn of the millennium is either outdated and has been discarded or is already included in the audit process.

For identifying the essential audit journals, the VHB partial rating list for accounting was used. This list includes the leading auditing journals and journals which combine technology, auditing, and accounting, such as the International Journal of Accounting Information Systems and Intelligent Systems in Accounting, Finance & Management. All journals which focus only on management accounting or finance were excluded from the search, as the research question focuses on external auditing. However, all accounting and internal auditing journals were included in the search, since publications on external auditing are often included in such journals. This yielded a list of 50 journals (see appendix for the list of reviewed journals).

Next, the electronic databases EBSCO, Web of Science, and Google Scholar were scanned. Focus was placed especially on advanced data analytics methods, as they are more unique to the auditing context than conventional ones. The following keywords were used in combination with the words "audit" and "assurance" in English and German:

- Big data
- Process mining
- Natural language processing
- Machine learning
- Data analy*(tics/sis)
- Artificial intelligence

Search alerts were set-up in the mentioned databases to make sure new publications were not missed during the writing of the dissertation.

The inclusion criterion for identifying relevant publications was that they dealt with data analytics methods applied to the audit context with real-world data. The emphasis on "applied" is because many papers discuss the possible use of data analytics but lack the actual application of the methods with real data (so called "thought papers").

4.1.2 Data analysis

Carrying out the outlined collection process resulted in an initial list of 260 records. These articles were then screened, based on the pre-defined inclusion criterion, to identify those of interest for the research question. The majority of the collected articles (215) are thought pieces on potential applications of data analytics in external auditing. Only 42 publications on using data analytics methods with real-world data for audit procedures were found[7]. These filtered articles were then thoroughly reviewed. The key results and insights are summarized in the following sections. The remaining publications were used as supporting literature throughout the dissertation.

[7] Three papers are not included in this count as an expert interview was conducted with one author. They are instead presented in chapter 6.

4.2 Overview of findings

The literature review yielded 42 publications dealing with data analytics methods applied to the audit context with real-world data. Table 7 divides these publications according to which of the three audit stages and audit areas they deal with.

Audit stage	Audit area	Count	%
1. Planning	Risk assessment	3	7.1
	Fraud	23	54.9
2. Performance	Internal controls	2	4.8
	Analytical procedures	3	7.1
3. Completion	Going concern	8	19.0
	Audit opinion	3	7.1
Total		**42**	**100.0**

Table 7 Illustration of use cases in literature based on the audit process

The majority of the publications focus on identifying fraud with data analytics methods, followed by methods to evaluate going concern assumptions. The remaining few publications focus on risk assessment, internal controls, analytical procedures, and audit opinion.

In the subsequent sections, the results and key insights of these 42 publications are presented, grouped by audit area. As some issues concern several publications, the critical discussion of the identified literature is bundled in the summary of this chapter.

4.2.1 Risk assessment

For risk assessment, three papers were identified, of which two focus on the detection of earnings manipulation and one on the prediction of material accounting misstatements.

Dikmen and Küçükkocaoğlu (2010) aimed to detect earnings manipulation with logistic regression and clustering. The results of the logistic regression achieved a 67% accuracy rate, which is probably too low to be used in a practical setting. Furthermore, the authors created four groups of risk clusters: very significant danger, serious risk zone, grey zone, and no evidence. This is an interesting approach and could give the auditor on the job more guidance for interpreting the results.

Dbouk and Zaarour (2017) also investigated earnings manipulation, this time using a Bayesian Naïve Classifier. The best result achieved was an accuracy rate of almost 87%. This sounds promising, but when looking at the confusion matrix, the algorithm misclassified four out of six earnings manipulation cases. If, following this pattern, 67% of earnings manipulation cases are not identified, it would probably not be useful for the auditor in real life.

Dechow, Ge, Larson, and Sloan (2011) created a logistic regression model to predict material accounting misstatements of US companies. In addition to the results showing an accuracy rate of only 61%, the model has a high number of false positives, which also limits practical application.

4.2.2 Fraud

The audit area of fraud is, by far, the most popular area for applying data analytics found in the literature review, with a total of 23 publications addressing this topic. Identifying fraud is like finding a “needle in a haystack”, because it is a relatively rare event (Hoogs et al., 2007). Furthermore, fraudsters actively attempt to hide their wrongdoing as best they can (Perols, 2011).

As a result, fraudulent attributes are similar to nonfraud attributes. This combination creates a challenge for any data analytics method. However, compared to the past, a significantly larger volume of data is now available for analysis, which can be both a curse and a blessing when it comes to identifying fraud (Tackett, 2013).

The respective accuracy rate of the reviewed methods is mentioned in the following summaries of the research publications, but must always be taken with a grain of salt. Unfortunately, a quantitative comparison of the results is not meaningful, given the diversity of approaches and methods employed. Some articles follow the best practice of splitting the datasets into training, validation, and testing set while others do not. The majority of publications used financial ratios as input variables, but the same ones are rarely employed. Some publications actively dealt with imbalanced datasets, while others neglected that aspect of the problem. Comparability limitations arise due to regional datasets from Greece (Kotsiantis & Kanellopoulos, 2008; Kotsiantis, Koumanakos, Tzelepis, & Tampakas, 2006a, 2006b), the United States (Debreceny & Gray, 2010) and China (Song, Hu, Du, & Sheng, 2014), each of which can be expected to contain local idiosyncrasies. Finally, the majority of papers effectively only distinguish between nonfraud and fraud cases, ignoring the considerable differences between various types of fraud (Koh & Low, 2004). Subsequently, the key findings of the fraud papers are briefly presented.

The results of Perols (2011) identified logistic regressions and support vector machines as more promising methods than decision trees and artificial neural networks. Song et al. (2014), who tried to identify fraud for Chinese companies, achieved the best results using a support vector machine with an overall accuracy rate of 89%. Kirkos, Spathis, and Manolopoulos (2007a) achieved even better results (90%) with Bayesian belief networks. For Dutta, Dutta, and Raahemi (2017), the best results were obtained by employing neural networks and decision trees. Katsis et al. (2012) tried to identify fraudulent financial statements with an ant colony optimization approach achieving an accuracy of almost 78% in their dataset. An advantage of this algorithm is that the rules applied by the Ant Miner are disclosed in "if-(and)-then" rule formats. Hoogs et al. (2007) experimented with a genetic algorithm approach, distinguishing between the various types of fraud and focusing on improper revenue recognition. In the end, the model accurately classified 63% of fraud cases and 95% of nonfraud cases correctly. A paper by Whiting, Hansen, McDonald, Albrecht, and Albrecht (2012) showed that the rule ensemble yields moderate levels of success in detecting fraud patterns in publicly available financial data.

That various methods yield the same outcome in a single regional setting is shown by Kotsiantis and Kanellopoulos (2008) and Kotsiantis et al. (2006a, 2006b) using data from Greece. In the Kotsiantis et al. (2006a)[8] publication, out of six classification methods, no single algorithm consistently outperformed the others. Therefore, a voting ensemble of all methods was created, which improved the accuracy rate to 91%. However, in two follow-up papers, two algorithms of the decision tree family alone achieved the same overall accuracy rate of 91% (Kotsiantis & Kanellopoulos, 2008; Kotsiantis et al., 2006b).

Perols, Bowen, Zimmermann, and Samba (2017) addressed the inherent problems of fraud (rare observations, missing explanatory variables, diversity of fraud) and analyzed the dataset with support vector machines. By adjusting the model to accommodate these factors, the performance of the previous model was improved.

Chen, Huang, and Kuo (2009) tested a mixed approach for identifying fraud in which auditors enriched the dataset with some external factors for the neural

---

[8] This paper investigated the detection of fraudulent firms and predictions of corporate bankruptcy. To avoid confusion and double counts, it was only counted as a "going concern" paper within this literature review.

network. With this additional feature, an accuracy rate of around 81% was achieved with a Type I (false positive) error rate of 16%.

Debreceny and Gray (2010) analyzed journal entries of 29 different organizations for fraud indicators by using Benford's Law, which is a well-known rule-based method. The analysis yielded the conclusion that either Benford's law does not apply to journal entries, or a very high number of red flags were raised. For the latter scenario, auditors would need to come up with a solution regarding how to analyze the red flags raised in the course of the audit.

Nigrini and Miller (2009) applied the second-order tests of Benford's Law to four transaction datasets: accounts payable, journal entries, annual revenue, and expense data. The second-order test expects that values ranked from smallest to largest closely approximate the digit frequencies of Benford's Law. The authors concluded that the second-order tests did detect errors in the dataset, rounding, and statistically generated random numbers, and therefore, might be a valuable addition to the auditor's toolset.

Singh, Lai, Vejvar, and Cheng (2019) focused on internal fraud and analyzed transactions from purchase processes using logistic regression. The model developed would not be implementable in practice, since the accuracy level was too low, and the false positives rate too high. Transactions from purchase processes were also used by Jans et al. (2010) as data. In this paper, a multivariate latent class clustering algorithm was applied, which has the advantage that an observation can belong to more than one cluster as expressed by probability. The data consists of technical information about the purchase orders (e.g., created by, when) as well as operational information about the purchase order (e.g., number of changes, changes after approvals/releases). The results revealed some interesting insights, especially for the case study company (e.g., mean of 25 changes per purchase order).

Abbasi et al. (2012) developed a meta-learning framework for detecting financial fraud, intending to outperform any previous results. Meta-learning accumulates experience from the performance of multiple base-learning methods (e.g., decision trees, neural networks) (Brazdil, Giraud-Carrier, Soares, & Vilalta, 2009). With the help of meta-learning, the inductive bias for each method is weakened. Their dataset consisted of industry-level and organizational context (e.g., comparison of performances), as well as financial ratio attributes. Quarterly and yearly financial statements were put into two distinct datasets, because empirical evidence has shown that different concealment methods are applied for the first three quarters of the year and the last quarter (Albrecht, Albrecht, & Albrecht, 2004). The authors compared the performance of the meta-learning framework with other research results (Cecchini, Aytug, Koehler, & Pathak, 2010; Gaganis, 2009; Kirkos et al., 2007a). The meta-fraud framework outperformed those results in the areas of accuracy and prediction of fraud and nonfraud.

Huang, Tsaih, and Lin (2014) tested the performance of growing hierarchical self-organizing maps, a type of unsupervised neural network. This method has the advantage of revealing the clustering pattern, which helps the auditor to understand the classification better. Patterns revealed, for instance, that companies with liquidity troubles and difficulties with paying short-term debts tend to record fictitious revenues. Such patterns are familiar to auditors.

Tackett (2013) investigated the usage of association rules to identify fraud. Association rules, also known as market basket analysis, are usually used to analyze consumer purchase behavior and what items are bought in combination. By looking into sales returns, the author was able to identify employees' fictitious credit card refunds.

Albrecht et al. (2004) argue that annual financial statement numbers are often too aggregated to reveal significant differences, which creates an opportunity for concealment methods. However, those concealment methods work better with annual data than with quarterly data (Abbasi et al., 2012). Abbasi et al. (2012) as well as Purda and Skillicorn (2015), therefore suggested including quarterly financial ratios in the analysis.

Goel and Gangolly (2012) applied natural language processing to financial communication materials, because "quantitative financial numbers contain redundant information that does not change when a company is committing fraud but the writing style and the presentation style employed to communicate financial information changes" (Goel & Gangolly, 2012, p. 76). The results of this analysis suggest that complex sentence structures (measured by ambiguity index), reading comprehension difficulty (measured by readability index), positive tone, passive voice, uncertainty markers, and adverbs can be linguistic cues associated with fraudulent financial statements (Goel & Gangolly, 2012).

Humpherys, Moffitt, Burns, Burgoon, and Felix (2011) also tried to decode deceptive language in fraudulent financial statements with linguistic credibility analysis. Both Naïve Bayes and C4.5 decision trees achieved an accuracy of 67.3%. The results revealed that fraudulent Management Discussion and Analysis (MD&A)[9] sections use significantly more active language (activation word ratio, imagery, pleasantness, modal verb ratio) and more words but have lower lexical diversity. The authors interpret the latter as managers trying to persuade readers of the veracity of the content by writing more while actually communicating less content. Purda and Skillicorn (2015) conducted a follow-up study of Humpherys et al. (2011) and achieved an even better accuracy of 82%.

According to ACFE (2019), fraud is typically committed within the company. Hence, Jans et al. (2011) investigated how process mining can be used to mitigate internal transaction fraud. A notable advantage of using process

---

[9] Part of public company's financial statements in the United States.

mining compared to classification methods is that the actual data is visualized, and no presuppositions in terms of output variables (fraudulent and nonfraudulent) need to be manually added. This makes the datasets for process mining objective and bias-free. The authors showed in their paper various methods to slice and dice datasets in order to discover any non-compliant behavior.

Thiprungsri and Vasarhelyi (2011) investigated how clustering analysis can be used for anomaly detection. They note that clustering has the advantage that researchers can avoid flagging the data manually with fraud or nonfraud markers. Having clusters with different characteristics is not necessarily a sign of anomalies, because legitimate explanations may exist for the variance. By performing a "test of one" for each cluster, these legitimate reasons can be checked quickly. For clusters where explanations are non-obvious or missing, further check-ups and audit procedures should be carried out.

4.2.3 Internal control audit

In total, two papers were identified, which focused on applying process mining to internal control audits. Both papers applied process mining because this method is uniquely suited to the requirements of this auditing task. That academia can play an essential role in the process of building useful audit tools is shown by process mining where the development of applications is driven by researchers (e.g., Günther & Van der Aalst, 2007; Van der Aalst, 2011, 2016; Van der Aalst & Weijters, 2005).

As shown by the use case of Zerbino, Aloini, Dulmin, and Mininno (2018), process mining can reveal the full process flows captured by an ERP-system. In this case, process mining showed that only 11% of the processes followed the internal control procedures. With such a low fitness rate, six significant process deviations were uncovered. Interestingly, two of the deviations were classified as “desirable,” because “they enhance the flexibility of the export process without compromising both its conformance and its operational execution” (Zerbino et al., 2018, p. 91). The remaining four significant process deviations posed a threat to the company from various perspectives (e.g., managerial, operational, or legal).

The paper of Werner and Gehrke (2015) showed that the multi-level-process-mining algorithm could be successfully applied to analyze the processes of various industries. This is a valuable result because it offers the promise of a universal tool that can be developed in the future and which does not require extensive tailoring to every use case. Within their paper, the authors only mention that the algorithm can be used for “process audits”. In the absence of a direct connection to the audit process, an application for the internal control audit is assumed, most likely for the creation of flow charts.

4.2.4 Analytical procedures

In total, three use cases performing analytical procedures were found in the literature.

Koskivaara (2000) used neural networks to predict monthly financial account balances. A generally applicable learning out of this paper is that training the model with more data yielded better predictions of monthly balances.

In a follow-up paper, Koskivaara and Back (2007) compared the performance of a neural network with the already implemented budgeting tool. On a higher level of analysis, when comparing overall revenue predictions, the neural network, on average, outperformed the traditional forecasting method. However, looking at each account prediction on an individual basis, the traditional budgeting method performed better. Both applications were designed for a continuous auditing environment. In the absence of this, both papers could be useful for the analytical procedure “trend analysis”.

Jans et al. (2014) carried out a study to investigate the use of process mining of event logs as an analytical procedure comparing actual process flows with planned flows (benchmarking). For this study, the authors analyzed the procurement data of a leading global bank. The analysis revealed payments made without approval, violations of segregation of duty, and deviations from internal defined procedures. Notably, none of these deviations had previously been detected by the internal audit team of the bank.

4.2.5 Going concern

In total, eight publications on using data analytics methods on the subject of going concern were identified. The evaluation of going concern, like the identification of fraud, has a binary outcome. Consequently, the majority of authors applied supervised machine learning methods.

Anandarajan et al. (2001) created three neural networks, which included qualitative indicators such as dividend reductions or debt defaults. Tested on the validation data, the best performer was a genetic neural network with an accuracy rate of about 95%. As the authors mention, the model provides a potential indicator for auditors. The final judgment call remains with the auditor, mainly because the model does not include variables such as future (turnaround) plans or management's leadership skill.

Kotsiantis et al. (2006a) used voting ensemble classifiers, which did not yield any practical results for the auditor as the accuracy of classifying bankrupt firms was very low (29%).

Martens, Baesens, Vanthienen, Bruynseels, and Willekens (2008) tried to predict the going concern assumption with support vector machines, rule-based classifiers, and ant-miners (ant colony optimization approach). Overall, the best results were achieved by logistic regression and support vector machines.

However, the authors argue that the results from the ant-miner method are more user-friendly in terms of interpretability.

The paper of Mendes, Cardoso, Mário, Martinez, and Ferreira (2014), in addition to the usual financial ratios, included a variable measuring accounting data inconsistencies of Brazilian health organizations. Examples are total liabilities and total equity that do not match total assets, or current assets and noncurrent assets that do not sum up to total assets. However, unsurprisingly, if the unaudited financial statements numbers do not add up, a self-fulfilling prophecy impacting the going concern assumption is reflected in the results.

Jiang and Jones (2018) used decision trees with gradient boosting, achieving a high accuracy rate of 95%. Before running the model, they conducted a relative variable importance (RVI) test, which indicated that market capitalization, annual market returns, and retained earnings to total assets are the strongest predictive values for the model.

Gepp, Kumar, and Bhattacharya (2010) used various decision tree algorithms and concluded that the best models depend on the specification of the misclassification costs, which are a subjective judgment. The papers of Lenard, Alrm, and Booth (2000), as well as Koh and Low (2004), did not use different data for the validation of the model. They both achieve accuracy rates exceeding 90%, but in this context, the high accuracy rate could well be the result of overfitting the model to the dataset, and no practical use case has yet been proven.

4.2.6 Audit opinion

The following three applications were identified in the academic literature investigating the prediction of the audit opinion.

Doumpos, Gaganis, and Pasiouras (2005) looked into the predictability of the audit opinion with three types of support vector machines, which resulted in similar accuracy rates. In a follow-up paper, Gaganis, Pasiouras, Spathis, and Zopounidis (2007) tested nearest neighbor, linear, and logistic regression in different settings. The authors experimented with combining financial and nonfinancial ratios as well as creating two industry-specific models. In terms of average classification accuracy, the nearest neighbor model outperformed the other two. This paper includes nonfinancial ratios, such as credit rating, which improved the accuracy considerably. Interestingly, creating industry-specific models did not always achieve better results.

Kirkos, Spathis, Nanopoulos, and Manolopoulos (2007b) compared the usability of decision trees, neural networks, and Bayesian belief networks for predicting the audit opinion. The best performance was achieved by the Bayesian belief network, resulting in an overall accuracy rate of 82%.

4.3 Summary

The majority of authors see the role of data analytics as a support tool for auditors. However, in "thought papers", where no application of data analytics with real-world data was attempted, there is frequently the expectation of benefits such as improvement of effectiveness, efficiency, and a higher degree of audit quality. Other examples regarding the expected benefits are broader, and range from becoming a business partner (e.g., Earley, 2015; Salijeni et al., 2019), via enabling the auditor to focus on critical items (e.g., Ruhnke, 2019), all the way to gaining a competitive advantage for audit companies (e.g., Omoteso, 2012).

Among the methods applied in the literature, within the time period analyzed, machine learning is the dominant approach, as can be seen in Figure 14. The underrepresentation of the conventional data analytics approaches of visualization, descriptive statistics, and rule-based approaches is probably because those approaches are well-known and have not benefited from the improved technological capabilities to the same degree as the more advanced data analytics approaches in the time period under review. If the timeframe was extended further into the past, these approaches might be more numerously represented.

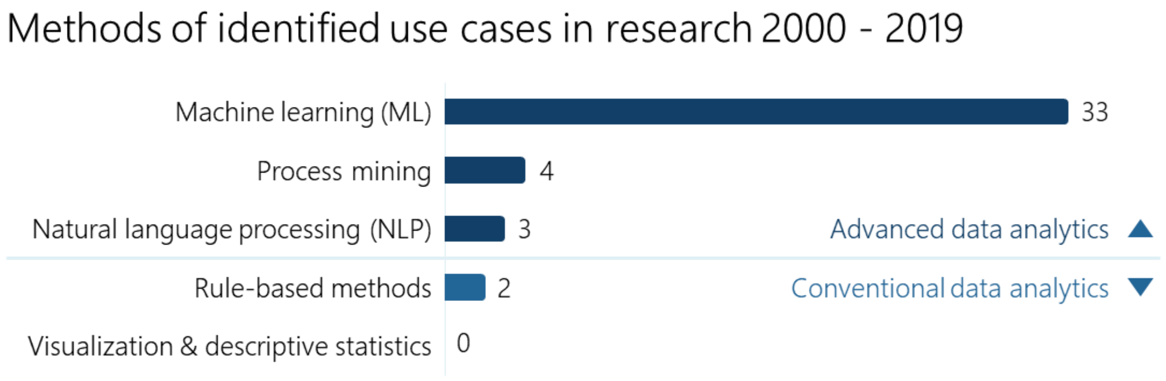

Figure 14 Methods of identified use cases

Among machine learning methods, supervised methods are more commonly applied than unsupervised methods. Since supervised methods require a labeled dataset, the challenge arises for practitioners to achieve reliable labels (Torgo, 2007), especially as the "common" audit data are not labeled by nature. Nevertheless, using labeled datasets allows the auditor to evaluate the results of the models in terms of accuracy and with a confusion matrix.

A recurring theme is that following up on false positives is tedious for an auditor, but not as bad as overlooking a real fraud-case (false negatives). The consequences of assigning an instance to the wrong class are known as misclassification costs, and for fraud cases, those costs are strongly asymmetric. False negatives are far more costly for an auditor than false positives (Abbasi et al., 2012; Weiss, 2004). When creating models, some authors included the misclassification costs. However, a value tag must be assigned to the misclassifications, which is subjective, and in the absence of scientific research, only estimated values can be used (Koh & Low, 2004).

Another general obstacle for using data analytics to identify fraud and going concern issues is the available datasets. Identified fraud and bankruptcy cases are vastly outnumbered by nonfraud or going concern instances, leading to imbalanced datasets, which are especially tricky for supervised methods. Consequently, strategies need to be used to deal with this imbalance. In addition, there is a possibility of "contaminated" data, meaning undetected fraud cases within the nonfraud data, as not all companies who "cook the books" are caught. Therefore, added labels must be viewed with caution. Additionally, there is no guarantee that the same kinds of fraud that occurred in the past will reoccur in the future. As supervised methods learn from the identified fraud datasets, it is questionable if unknown fraud patterns will be detected.

Another barrier is that while such methods can give a comprehensive indicator for fraud, they do not reveal the source of tampering. In practice, such an indication could create a potential bias for the auditor, and be the catalyst for a prolonged audit because the search for the needle in the haystack begins. Especially looking at the results can create a potential bias for the auditor. Take an example where a model tells the auditor that there is a 99% probability of the presence of fraud, but the auditor does not find conclusive evidence. How should an auditor react in such a situation? Can he ever stop looking for the needle in the haystack without neglecting his professional duty? The current ISA standard does not give guidance, as using machine learning within auditing is not yet addressed therein. Without sufficient and appropriate audit evidence at hand, the auditor has no actual proof of fraud. This makes it clear that performing data analytics is only an analysis, and further audit procedures to provide the concluding audit evidence are still necessary.

The majority of papers use financial ratios. The limited performance of the applications presented suggests that the financial measures and classification methods employed were insufficient (Abbasi et al., 2012). Once the financial ratios incorporated in the model are general knowledge, it will be easier for fraudulent companies to manipulate their financial statements results to provide the right financial ratios without being detected.

Finally, models that solely employ quantitative data can lack critical qualitative features. This leads to a new inherent risk. Therefore, blindly trusting high-accuracy results is not sufficient. Even with new powerful tools, the auditor's professional skepticism and judgment will still be required and must be applied with vigor.

Figure 15 summarizes the use cases identified in the literature review.

In conclusion, although literature exists, various limitations apply to it. One thing it does make clear is that there is no single "magic algorithm" for auditors at the moment (Abbasi et al., 2012; Hoogs et al., 2007). The real test of the usefulness of data analytics is provided by its use in practical contexts in the daily life of auditors. This is the subject of the next chapter.

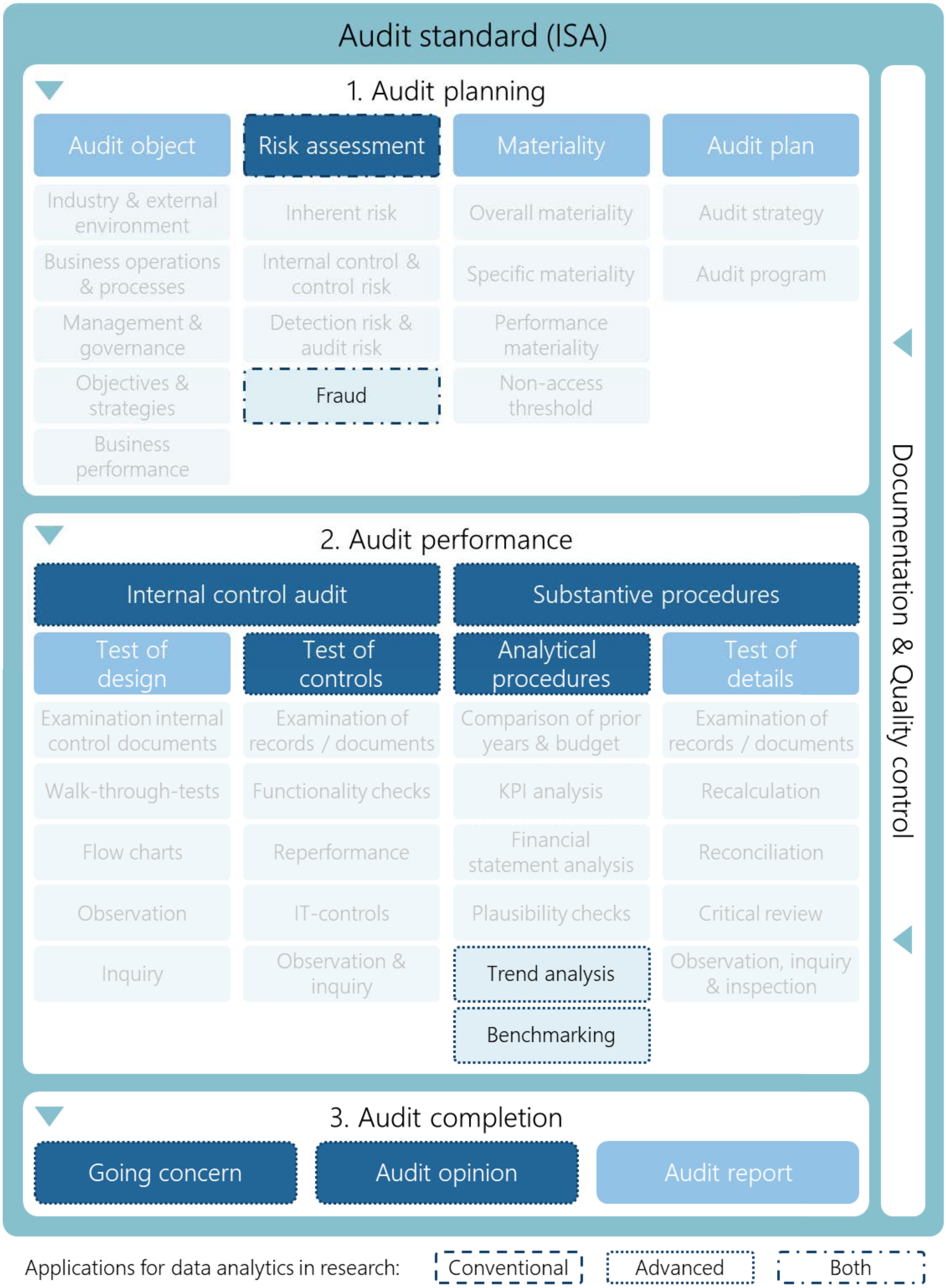

Figure 15 Illustration of applications identified in research

5 Current applications in audit practice

5.1 Research methodology

This chapter is devoted to the second research question:

Which data analytics methods have been implemented in the external audit process?

Given an identified lack of information on current applications of data analytics in audit companies (Appelbaum et al., 2018; Earley, 2015), interviews with eight audit companies were conducted to address this research gap. These interviews provide qualitative insights into the subject and can serve as a roadmap for future quantitative research (Wilson, 2009). In-person interviews can be particularly useful when examining the use of technology (King, Cassell, & Symon, 1994). They enable the researcher to follow-up on spoken content immediately and to clarify uncertainties (Hair, 2011).

Like every research method, interviews have advantages and disadvantages. A strength of interviews is the flexible approach to knowledge discovery, which can provide insights and enhance findings with context (Hair, 2011; Robson, 2002). At the same time, collecting data through interviews raises concerns about reliability due to the lack of standardization. This drawback was mitigated by the creation of a topic guide, which allowed the author to follow a consistent protocol in every interview (Yin, 2009). Another disadvantage of interview data is potential bias, both in the selection of interview partners and in the communication process itself. To deal with this issue, published sources were used to validate the interview results, and to mitigate bias within the collected data. To mitigate selection bias, interview partners were chosen based on their involvement in using data analytics in the auditing process (Gläser & Laudel, 2010).

5.1.1 Sample selection

According to Wilson (2009), drawing a random sample as performed in quantitative studies is neither necessary nor realistic for qualitative research. However, the selection process must be explained and justified. For the selection process, the following two criteria were defined.

First, the selection focused on audit companies belonging to a network. The reasoning is that this segment of the audit industry has more resources for data analytics and higher purchasing power for new tools compared with smaller, often regionally focused, auditing companies. Based on this selection criterion, the top ten auditing firms in Europe by revenue ranking (International Accounting Bulletin, 2017) were used as a starting point. The list contains the following audit companies in alphabetical order: Baker Tilly, BDO, Deloitte, EY, Grant Thornton, KPMG, Mazars, Nexia International, PwC, and RSM.

Second, the selection criteria were narrowed to the Swiss branches of these networks for the following two reasons. First, Switzerland is a country with one of the highest labor costs in Europe (Federal Statistical Office, 2018). Under the existing assumption of realizing efficiency gains with data analytics (Appelbaum et al., 2018; Issa et al., 2016; Salijeni et al., 2019), strong incentives exist for its early implementation in Switzerland. Furthermore, Switzerland has a skilled workforce to support implementations of data analytics in auditing (World Economic Forum, 2019). Given these characteristics, Swiss audit firms are well suited for answering this research question, as they can be expected to be early adopters of new and innovative audit technologies.

Requests for interviews were sent by e-mail to partners, directors, and managers actively involved in auditing. Eight companies replied positively and agreed to participate. One of the smaller audit companies declined because the subject is still new for them and they are only just starting to think about this topic. A second company declined due to a lack of personnel capacity.

Interview appointments were scheduled for 60 minutes, because “anything under half an hour is unlikely to be valuable; anything going much over an hour may be making unreasonable demands on busy interviewees, and could have the effect of reducing the number of persons willing to participate” (Robson, 2002, p. 273).

5.1.2 Data collection

The objective of the data collection is to gather enough relevant data to understand current applications of data analytics in the audit process. This also serves as the basis for identifying implications and future applications later.

The topic guide contained twelve questions, which is within the range of eight to 15 questions for an hour-long interview recommended by Gläser and Laudel (2010). The questions of the topic guide were derived from the research question and drawn from calls for research in the academic papers analyzed during the literature review (Alles & Gray, 2016; Alles, 2015; Appelbaum et al., 2017; Earley, 2015; Gepp, O'Neill, Linnenluecke, & Smith, 2018; Issa et al., 2016; Ruhnke, 2019; Salijeni et al., 2019; Titera, 2013). The complete topic guide can be reviewed in the appendix (“Topic guide for audit companies”).

To ensure all interviewees approached the topic from a shared understanding of “data analytics”, the definition used in chapter 3 was provided at the beginning of each interview, and the different methods of data analytics were presented with an example for each method.

The time dimension for data gathering was “point in time”, also known as a “cross-sectional study” (Blumberg, Cooper, & Schindler, 2008). All interviews were audio-taped, with the explicit permission of the interviewees, to get a complete copy of the spoken content. Recording, rather than transcribing the content on the spot, allowed the interviewer to concentrate on the discussion itself (Robson, 2002). All interviews were carried out in the interviewees' native language (Swiss German or German) to avoid language barriers.

An obstacle for audit companies to participate was the confidentiality of information regarding new innovations and competitive advantage considerations. All interview partners were assured anonymity to address this concern.

5.1.3 Data analysis

As mentioned before, in total eight interviews were conducted. Through the interviews conducted in the data collection stage, a large amount of data was collected. To answer the research question adequately, the main focus lies on critically evaluating the information provided rather than reproducing the spoken content word by word. By comparing the gathered statements, evaluating them and bringing them into context, a holistic view of this topic is achieved. In an exploratory setting like this, coding would not be a suitable method, according to Meuser and Nagel (1991). Therefore, a thematic transcription was done to condense the accumulated data to the key arguments relevant to the research question. Alvesson (2011, p. 59) defines thematic transcription as summarizing "the major content of what is being said".

Once the transcription of all interviews was completed, the records were re-checked to ensure no important detail was missed during transcription. The collected data was thematically clustered into the following categories: methods used, tools applied, audit stages concerned, motivation of use, challenges, and audit standard. Answers were allocated to the respective segment.

Out of the eight interviews conducted, five participants are currently either focused on data analytics or technological audits and had a generally positive attitude towards data analytics, while three had a more critical attitude. In the following subchapters, the provided answers were verified and extended with corresponding sources from literature where possible (Hair, 2011). Additionally, the annual reports of the Swiss Federal Audit Oversight Authority (FAOA) were also used to verify the content. Since 2016, the FAOA devotes a chapter of its annual report to the current use of data analytics.

The results are presented without anonymized abbreviations. This is for the following two reasons: first, to protect the anonymity of the interviewees; second, since the research objective is to get a better understanding of current practice, and thus to assemble the combined experience of the interviewees, rather than focus on the individual or the respective companies.

5.2 Overview of findings

Five interviewees acknowledge an increase in the popularity of the topic. Six auditors feel pressure from their clients to perform data analytics or to at least mention data analytics in their tender offers. One participant sensed pressure from the FAOA when auditing larger companies. Three interviewees had the impression that using data analytics pays off for all sorts of customers.

In the next sections, the data analytics methods are addressed first (5.2.1). As the methods are incorporated in data analytics tools, the subsequent subchapter is devoted to the tools in place (5.2.2). This is followed by an explanation of where these methods and tools are applied within the audit process (5.2.3). Since the use of data analytics is currently not required by the standard setter, a subchapter is devoted to the motivation of the auditors using data analytics (5.2.4). Furthermore, auditors face various challenges in applying data analytics methods; these hurdles to implementation are addressed in a separate section (5.2.5). The last subchapter focuses on how the standard setter influences the use of data analytics in current practice and provides a summary of all applications, organized according to the audit process (5.2.6).

5.2.1 Methods

Figure 16 shows which data analytics methods are currently in use, based on the eight interviews.

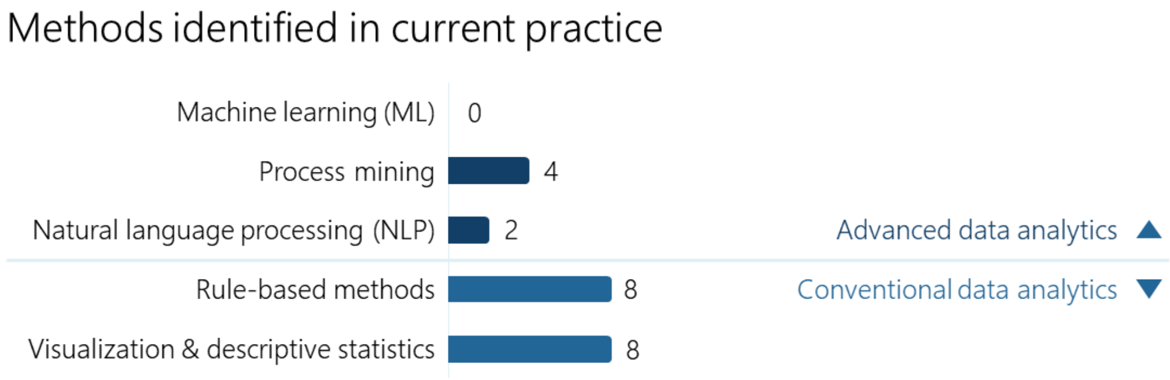

Figure 16 Identified methods in current practice

When asked which data analytics methods have been implemented in the audit process, all participants mentioned rule-based data selection as well as descriptive statistics and visualization. For the prevailing data analytics methods currently in use, human judgment is needed to interpret and check the results for plausibility. The results are easily interpretable for non-IT savvy people, according to the majority of the interviewees. Nevertheless, the auditor cannot blindly trust the rule-based methods, because tests performed are only as good as the rule defined. Therefore, it is crucial that the auditor knows and understands the calibration of the rules.

Four interviewees mentioned using process mining. Within an audit context, this method is primarily used for internal controls. To analyze them, event logs are needed, which currently can either be found in an ERP-system or need to be created by combining the data of various systems. It is crucial that the auditor pays attention to whether the full process is reflected in the ERP-system. For any control outside the system, process mining cannot be applied, as the required event logs will not be available. With the help of process mining, the auditor and audit client gain transparency on how processes flow within the ERP-system.

Two participants mentioned the use of natural language processing to capture textual data, extract concepts, and verify the content of journal ledgers or account balances with documents. To use natural language processing, the available documents must be machine-readable. To fulfill this prerequisite, optical character recognition (OCR) technology is often used, which can turn scanned images into a text file format. According to one of the participants, the advantage of natural language processing is that it leads directly to evidence or identifies differences, whereas with other data analytics methods, only an analysis is performed, which could require further follow-up audit procedures.

No interviewee explicitly mentioned the use of machine learning in auditing. This finding is in line with Westermann and Spindler (2017), who claimed that machine learning is barely a topic in auditing practice today. However, it should be noted that both process mining and natural language processing make use of machine learning algorithms.

5.2.2 Tools

How data analytics methods are implemented varies by audit company. For performing rule-based methods, descriptive statistics or visualization, spreadsheet tools such as Excel are widely used. Excel allows the auditor to use pre-programmed macros, create pivot-tables, generate various diagrams and analyze data with formulas (e.g., minimum, maximum, average, duplicates). Using spreadsheet tools helps the auditor to slice and dice the data as needed, get more familiar with the data itself, and gain a better understanding of it (Küng, Hamberger, & Schweitzer, 2017).

In combination with visualization and descriptive statistics, dashboard tools are being used. As mentioned, an advantage of such dashboards is the drill-down functionality, which allows the user to go directly to the represented journal entries. During the interviews, one example mentioned was self-built SQL (structured query language) queries with rule-based tests.

For tools more advanced than Excel, audit companies are confronted with a make-or-buy decision. Two participants explicitly mentioned that the "make" decision is a privilege of the biggest four or five companies, as smaller companies lack the financial, technical, and human resources. However, even for smaller companies, solutions exist in the marketplace and are available as a "buy" option. Even if dedicated human and technical resources are limited in non-Big 4 companies, service providers exist that perform data analytics for them. Such services include the performance of the data analysis for a specific audit objective, along with complete documentation of the steps performed. In addition, as all interviewees belong to an international network of companies, two interviewees from non-Big 4 companies revealed that they also depend on the investments of their international networks or on cooperation within the network to develop such tools.

Four interviewees said that data analytics tools developed in-house were available. Three participants mentioned using products such as CaseWare IDEA (Interactive data extraction and analysis) or ACL (Audit command language) for data analytics during audits. The advantages mentioned for CaseWare IDEA were a pre-built library with various rule-based tests and the possibility of creating own rule-based queries. The software also allows the user to perform descriptive statistics and visualization. Especially useful in the audit context is the automatic documentation of every performed procedure, resulting in a traceable audit trail. One interviewee further mentioned using a spin-off tool of IDEA, which can be integrated into Excel as an add-on.

Four interviewees also explained that they have industry-specific or even company-specific data analytics tools. The same interviewees revealed that they hire data scientists to perform data analytics.

The existing options regarding process mining are an open-source platform called ProM, and proprietary commercial solutions such as Celonis or Enterprise Discovery Suite. The latter two are more user-friendly, but have fewer features than the open-source platform ProM, which benefits from 1500 plug-ins created and maintained by the community (Van der Aalst, 2016). Even though ProM offers more functionality than the commercial solutions, expert knowledge is required to implement all the available options. Three out of four interviewees who use process mining mentioned that they use a commercial provider.

Two out of the total of eight interviewees have a data analytics tool customized for clients using SAP, which contains several rule-based tests, especially concerning the internal controls and the usage of SAP. This tool also offers visualization capabilities and descriptive statistics. Additionally, this tool allows the auditor to compare the usage of SAP by the audited company with other companies, which are anonymized in the system (benchmarking).

Of note was that some implemented tools originally had their roots in the consulting departments of the audit companies, but are now used when needed in the course of an audit.

In current practice, such tools are commonly subsumed under the terminology Computer Assisted Audit Techniques (CAATs) (e.g. Alles, Kogan, Vasarhelyi, & Brennan, 2006; Earley, 2015; Wang & Cuthbertson, 2014). When looking at the tools and methods currently in use, the finding suggests that audit companies are still at the beginning of using advanced data analytics. Furthermore, various tools are used to perform data analytics. Half of the interviewees were of the opinion that, from a technological perspective, more methods and tools could be applied to auditing than are actually used today. The potential is there, but practical implementation in audit practice of more advanced tools is yet to come.

5.2.3 Audit stages

Three out of eight interviewees hold the opinion that data analytics can be used in every audit stage. In addition, one interviewee observed that all audit procedures that include the word "analysis" are ripe for data analytics applications.

The following subchapters go into detail on where the methods and tools are currently used within the audit process. The findings are presented following the audit process outlined in chapter 2.

5.2.3.1 *Audit planning*

When asked in which stage of the audit process data analytics tools are used, all participants mentioned risk assessment. Conventional data analytics methods or the tools in place help, for example, to identify the top ten clients of the audit client, perform margin analysis or determine the distribution of revenue.

As no guidance exists on how to apply process mining within an audit engagement, three different fields of application have been observed. The first area where process mining is applied is during risk assessment. Applying it in this early stage minimizes the risk of uncovering unexpected process flows, which could prolong the audit performance stage. However, given the possibly visual outcome of process mining (see Figure 17), it is understandable that auditors are tempted to choose a more cautious approach.

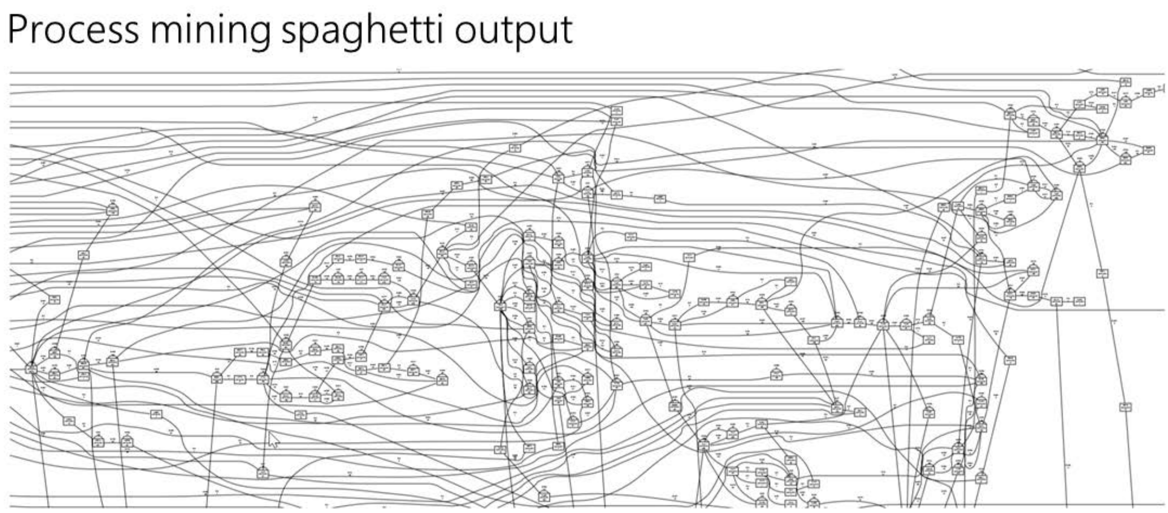

Figure 17 Example of "Spaghetti" process retrieved from Günther and Van der Aalst (2007, p. 330)

All participants use data analytics to perform mandatory journal entry testing to identify fraudulent activities or material misstatements[10]. The participants use methods such as rule-based tests (e.g., SQL-queries), visualization, and descriptive statistics to perform the testing. Rule-based queries specifically search for entries on weekends or holidays, amounts just below approval limits, missing or double entry numbers, abnormally high or low numbers. Some data analytics tools have already incorporated rule-based methods and offer the option of flagging journal entries with low, medium, or high-risk labels. Once the tools or methods are applied, it remains the duty of the auditor to rationalize, interpret, and judge the results of the analysis.

Four participants reported using data analytics to analyze the audit object, especially the business performance. Two interviewees even mentioned that external data is purchased to evaluate the performance of the audit clients against the performance of peer-groups or competitors. In addition, for some audit clients, business performance can be partially validated with

[10] The risk assessment of fraud takes place in the audit planning stage. However, the actual journal entry testing is conducted in the audit performance stage. For consistency with the presented audit process, it is covered in this subchapter.

macroeconomic data. This helps the auditor to get a better feeling for the industry and external environment of the audit client.

Side effects of the tools used, such as suggesting materiality thresholds or guiding the audit team through the audit, were also named as benefits of applying such tools.

Two participants perceived an expansion of the audit process insofar as the data export becomes the first step even before the audit planning stage starts. The majority of interviewees noted that by using data analytics, more audit work is pushed into the audit planning phase. This is because the data is primarily analyzed to yield early indications of risky audit fields and not for producing audit evidence. That, in turn, goes hand in hand with ISA 315.5, which states that any insights gained from risk assessment procedures do not provide sufficient appropriate audit evidence for forming an audit opinion. Consequently, for any identified risks in the audit planning stage, the required audit evidence still needs to be collected in later audit procedures. By applying data analytics already in the audit planning phase, the auditor is also protected against running into unpleasant surprises late in the process, which could prolong the audit and cause the agreed deadline for the audit report to be missed.

Using data analytics for risk assessment in the audit planning stage indicates which fields warrant further attention. Those areas that have been identified for further review can be directly included in the audit plan. Three participants mentioned that the use of data analytics during the performance stage needs to be actively incorporated in the audit plan, especially as getting the data requires coordination with the audit client, and it needs to be ensured that the data is available on schedule. Frequently, the client needs lead time to prepare the data or grant access to the technical systems. If the data is not ready when requested, the audit cannot be performed as planned, which can create delays or forces the auditor to re-plan the audit performance stage.

5.2.3.2 *Audit performance*

For the audit performance stage, all methods and tools named by participants for the internal control audit and substantive procedures are discussed.

One participant suggested that process mining has the potential to eliminate internal control audits if the complete process is fully captured in the ERP-system. However, he clarified that this is a provocative view and the audit community is divided on this approach. Two further interviewees were more reserved about the application of process mining for internal control testing and considered it rather as a support for walk-through tests or as an aid for drawing flow charts within the test of design. They also saw process mining as useful for the test of controls in general.

Around 63% of the participants use rule-based queries to test the IT-controls. Examples given by the interviewees included testing the separation of functions, checking access rights, or performing the three-way-match, which is the reconciliation of purchase order, goods receipt, and invoice. For one participant, process mining is used as a substantive analytical procedure, which is a function supported by Jans et al. (2014). As actual processes are compared with target processes it is seen as a form of "benchmarking". The application of the SAP tool mentioned in 5.2.2 also falls into the benchmarking category.

All participants reported using descriptive statistics and visualization for substantive analytical procedures such as comparison of the previous year and budget, KPI analysis, trend analysis, analysis of the financial ratios, plausibility checks, and benchmarking. If external data was purchased in the audit planning stage, this data can be reused for benchmarking. Some data analytics tools, such as CaseWare IDEA, have a built-in function for various sample selection methods, which can be applied during the test of details.

Two interviewees reported applying natural language processing for the audit tasks "examination of records and documents" and "reconciliation of information". The example given for the latter audit task was the comparison of disclosure checklists with financial statement notes. In addition, natural language processing is used for the reconciliation of confirmation letters (e.g., bank confirmation letters to verify the bank balance as of a specific date).

One interviewee shared a success story of discovering material misstatements thanks to applying data analytics during the audit.

5.2.3.3 *Audit completion, documentation, and quality control*

Only one participant mentioned using data analytics for analytical procedures in this stage. Three participants suggested that a better basis for determining the opinion is provided when data analytics was used during the previous audit stages.

No use cases were identified concerning quality control and documentation. However, some existing data analytic tools (e.g., IDEA) have an add-on feature that automatically records each audit step performed within the tool.

At the end of a data analytics driven audit (subsequently called data driven audit), the auditor is still in charge of deriving an audit opinion. One participant expressed an appreciation for including data analytics in external auditing, claiming it enables the audit profession to reclaim its professional and skeptical judgment. The participant felt that over the last several years, too much focus has been given to filling out endless checklists, to the detriment of critical thinking.

5.2.4 Motivation

One motivation for implementing data analytics tools in the audit process, named by all participants, was effectiveness in terms of higher audit quality. Three of the participants understood this as gaining a better understanding of the audit client, while four saw it as resulting from the capacity to perform audits upon the whole populations instead of only samples.

Additionally, five participants expected efficiency gains from being able to perform the audit faster. Three of those five expected cost reductions from realizing those efficiency gains. However, in the short term, it is doubted if cost reductions are achievable since audit companies would have to invest either in the licenses for purchasing the data analytics tools, or in developing and maintaining such tools. This argument will be revisited in the next subchapter, "Challenges".

Other motivations to use data analytics are to be perceived as a valued business partner for the audit client because new insights are provided, according to four interviewees. Simultaneously, the service of providing new insights was criticized by three interviewees who see such add-on services as beyond the responsibility and primary goal of an auditor. They claimed the auditor's responsibility and primary job is to judge the conformity of the audited financial statements with the applied accounting standard in all material matters, and not to bring new insights.

It was pointed out that with such innovations and the related insights, the line between auditing and advisory is blurred. In the case of process mining, the audit client gets an audit opinion and a transparent evaluation of the company's processes. However, only providing the results of data analytics is not advisory and the independence of the auditor by providing those insights, which were produced during the audit, and not handing out recommendations does not break current regulations. When such insights lead to an advisory project, auditors find themselves in a potential grey area depending on the audit client (e.g., public-interest entities) and current regulations. The European Union explicitly prohibits audit firms from undertaking advisory projects for public-interest entities, which they audit as well, when it comes to designing and implementing internal controls related to the preparation or control of financial services (e.g., regulation European Union Number 537/2014). Similar regulations do not exist in Switzerland to date. However, providing advisory services first and auditing their own services later is not allowed by the rules of EXPERTsuisse (2015).

The interviewees were asked whether they expected audit fees to increase when providing new insights with data analytics. All responded that, on the contrary, clients assume fees will decrease, and most suggested that this was due to the automation that data analytics tools are expected to allow. Add-ons such as process mining are expected by some audit clients to come at little or no additional cost, as they think that the auditor will have benefitted through efficiency gains. However, five participants commented on the "set-up costs" incurred when applying data analytics. Two participants mentioned that it can take two to three audit cycles per audit client until these costs are amortized. This indicates the existence of a clear gap between client expectations and the reality of (at least initially) increased costs incurred by auditors in the implementation of these new tools.

5.2.5 Challenges

All interviewees reported current challenges when applying data analytics in auditing. These challenges are presented in the following subchapters, separately for audit clients and audit companies.

5.2.5.1 *Challenges with audit clients*

The biggest hurdles for the implementation of data analytics by the auditor lie on the client's side, according to the interviewees. Various problems were observed regarding data extraction, data quality, its respective assessment, and numerous data models, to name just a few.

Three non-Big 4 participants revealed that audit clients from the small and medium business sector especially dislike or distrust using data analytics in auditing. The reasons include data security, missing expertise in terms of getting the needed information out of their systems in the right format, fear of auditors changing the system or simply a desire to stick to the known audit process. Concerns regarding data security were frequently mentioned, as plenty of personal data is stored in the clients' systems. Clients express concerns about whether the audit companies can appropriately safeguard the data.

Additionally, some industries (e.g., banking, insurance) do not want to provide this data to external parties and instead set up a computer for the auditor to access data on the client's premises. This means that the data analytics tools must be usable on the computers of these companies if they are to be applied at all. As some audit companies use third-party providers (domestic or foreign), the auditor must receive the client's approval. Further critical concerns arise in cases where the client's data is analyzed outside the country in data centers. Given a recent tightening of data protection and privacy laws, there is an increased awareness of and sensitivity towards these issues throughout the business community.

Even if the client is, in general, receptive to a data driven audit, the next question is whether the needed data is available and of the required quality. Even if the data exists in the client's systems, it needs to be extracted first. All participants mentioned difficulties in getting the required data from clients. As multilayered approval processes are in place on the client's side to safeguard the data, just getting the permissions for the data extraction can be a time-consuming task. In case of insufficient knowledge of how to export the data on the client's side, audit companies assign their IT-auditors or data scientists to conduct this task.

After dealing with the data extraction and data security issues, the auditor has to make sure that the data meets the quality requirements (mentioned by five participants). Some audit companies have IT-auditors or data scientists who check data quality for consistency and completeness.

The next step is the preprocessing phase, which can again be time-intensive, according to five participants, especially if a diverse IT-landscape, media disruptions, and various data formats exist. The biggest problem for three participants is the missing common data model and variety among data lakes (data repositories). The handling of such heterogeneous data currently requires a lot of time for data preparation, which is prone to errors if done manually. The heterogeneous environment of data standards and models is a concern, especially when auditing clients without an ERP-system. Since there is no single dominant accounting software provider in Switzerland (unlike ERP-system providers, where SAP is the current market leader), various data models exist, and not all necessary information is captured by the system (e.g., user changes in the system).

In the end, it is the audit client who decides which audit model – traditional or data driven – he wants, because he chooses which audit firm is appointed as auditor, and both audit forms result in the same outcome (audit report). The interviews conducted for this dissertation revealed that currently there are two segments of clients. On one side is the "price-sensitive client", who is interested in low audit fees and receiving an unqualified audit report. On the other side are clients who are looking for new insights about their company ("insight clients"). The implications for both client groups and their impact on implementation of data analytics in external auditing are discussed in subchapter 7.2.2.

5.2.5.2 *Challenges within the audit companies*

The challenges of using data analytics in the course of an audit are not limited to audit clients, but also affect the audit companies and their employees.

As stated by the participants during the interviews, having the right tools is a good start, but without the know-how of the people who use them, the best tools are worthless. For more advanced data analytics tools, data scientists or IT-auditors are needed to ensure adequate application. For tools applied directly by the "regular" auditor, internal and external trainings are in place at most of the interviewees' companies.

Regardless of who carries out the analytical work, in the end, it is still the auditor's task to interpret and classify the results. Two participants mentioned that it can be challenging to interpret the results correctly, translate them into a business context and turn them into actionable audit recommendations that audit clients can understand. One participant mentioned that data analytics could tempt some auditors to jump too quickly to a conclusion which is not actually supported by the data at hand.

Besides the training requirements, five participants pointed out that the actual application of the developed methods and tools depends on the audit team's attitude towards data analytics, especially the attitude of the engagement leader. Various negative emotions in connection with data analytics are held among employees, which can be also found in literature (e.g., Bartmann, Hufgard, Streller, & Weltner, 2018; Goldshteyn, Gabriel, & Thelen, 2013). Participants named a lack of trust and the auditor's lack of knowing how to handle data analytics tools as reasons to avoid using them. A lack of understanding of data analytics tools leads to a lack of confidence in using them and sabotages successful applications. Another obstacle identified in literature (see also Alles et al., 2006) is the potential for information overload. When using tools such as process mining with full transparency of the processes conducted, the auditor could be overwhelmed by the number of exceptions, far beyond their processing capabilities.

Another hesitation noticed during the interviews in regards to using data analytics tools is the deviation from the audit process which the auditor is used to. Comments such as “we have always done it this way; why is this not good enough anymore?” commonly arise. This leads to an uneasy and reluctant feeling for some auditors when performing data analytics. As two participants suggested, trust in the technology can be increased by performing additional audit procedures to verify the evidence derived from data analytics. However, performing data analytics and additional confirmatory audit procedures increases both effort and costs, which in turn lowers profits for the audit engagement. Nonetheless, at least initially, such dual-performances may be necessary to establish the needed trust within the workforce.

As half of the participants mentioned, applying data analytics is a cost-benefit tradeoff. When questioned what determines this evaluation, no consensus emerged from which it would be possible to arrive at the client characteristics beneficial to the use of data analytics.

Especially in the interviews with larger audit companies the term “change management” was used, indicating that an active transformation is taking place. To support this transformation process, various active nudging strategies are in place to make sure the tools are used. Two participants, who are responsible for implementing tools, concluded that simply notifying staff that the tools are available and trusting that everyone will use them voluntarily does not work. Without constant nudging, follow-ups, and tracking, a thorough implementation does not take place.

As the application of data analytics is still at an early stage, it is too early for investment calculations or tracking returns on investment, especially as some tools are also used by the advisory departments. However, measures are in place to track the usage of the tools implemented in auditing based on the information from two participants.

Three out of eight participants feel that it takes some courage to apply data analytics and to refrain from performing any additional other audit procedures covering the same requirements instead. They suggested that guidance from the standard setter would help overcome these hesitations.

5.2.6 Audit standard

Five out of eight participants requested guidance from the standard setter, while the remaining interviewees said the available standards only need to be reinterpreted, and that changing them is not necessary. In any case, four interviewed participants said that the audit standard plays a secondary role when applying data analytics in auditing. Five audit companies developed internal guidance to address the gap created by the missing guidance from the standard setter. For two participants, the biggest challenge is ensuring that the data analytics tools in use comply with the current standard, as the standard does not mention the use of such tools.

In mid-2015, the IAASB launched a data analytics group, which released a request for input in September 2016. In this paper, the working group states the "ISAs do not prohibit, nor stimulate, the use of data analytics" (IAASB, 2016, p. 8), but also concedes that the current standard was written in a different era. The standard setter did not and could not anticipate the technological advancements in auditing at the time of writing them (IAASB, 2016).

The replies to their request for input exemplify the difficult situation of the IAASB. On one hand, guidance – even non-authoritative guidance – is requested. On the other hand, loud voices proclaim the "IAASB must be careful not to prematurely commence standard-setting activities related to data analytics, especially if doing so could have unintended consequences – such as restricting innovation." (IAASB, 2016, p. 19).

When questioned about upcoming regulations, especially non-Big 4 companies were afraid of standards only the Big 4 companies can comply with. All non-Big 4 companies agreed that if a data analytics standard is issued, it must be manageable also for smaller audit companies. Otherwise, barriers to entry into the audit market are created, and smaller audit companies may be regulated out of business. However, the working group points out that even for smaller audit companies, software solutions are offered by the market (IAASB, 2016). Services for performing data analytics for auditors can also be found in the market.

Compared to other audit standard setters (e.g., United States Auditing Standard Board, which is part of the American Institute of Certified Public Accounts), the IAASB has the particular challenge to provide regulation for many countries and across continents around the world.

There are indications that the IAASB will gradually incorporate the use of data analytics in the current revision of individual standards. As announced in May 2019, the revised risk assessment standard (ISA 315), the quality management standards at a firm level (ISQM 1), and at an engagement level (ISA 220) address the use of automated tools and techniques (IAASB, 2019b). As the IAASB is a principle-based standard, it can be assumed that the standard body will stick to those roots (Schneider & Bradtke, 2017).

One area in which it appears that guidance from the standard setter would be useful, is in how to deal with the increased number of anomalies found when auditing larger samples. When whole populations containing thousands or millions of journal entries are audited, it is inevitable that more "red flags" appear, according to three participants. Current literature (e.g., Salijeni et al.,

2019) often raises the question of how to handle this phenomenon. Although not directly related to full population audits, the current ISA 530 "audit sampling" uses the terms "tolerable misstatement" and "tolerable rate of deviation", which both offer some room for flexibility.

When this topic was raised with the participants, some noted that having an increased number of findings increases the dialogue with the client. Beyond this, they mentioned different approaches to dealing with the challenges raised. The strategies are frequently specific to the data analytics methods employed. Some noted that the identified exceptions can often be quantified with a monetary value that allows the auditor to determine if they are material. Another approach was to cluster the identified exceptions and to investigate these clusters further. Some clusters might be the result of a flawed set-up in the system, some might be human mistakes, and some could also lead to the identification of fraud.

It is not clear whether a full population audit performed with data analytics conforms to a full population audit as defined in the current ISA standard. First, the auditor needs to make sure that all data is captured in the system to perform an analysis of the whole population. This is especially uncertain if internal controls are performed manually. Even if all data is available, it is questionable if the "full population" audit is performed as the human auditor does not manually review the full dataset, but merely follows up on identified red flags. This can be countered with the argument that the machine "reviewed" the entire population and that the machine is simply an extension of the human auditor, carrying out what the human auditor told it to do. Even with conventional data analytics methods, tools are used by the auditor – but it would never be concluded that the spreadsheet did the audit. There is no clear boundary here, but at some point, the standard setter will need to address this question.

Besides the fundamental discussion about the meaning of a full population audit, two interviewees questioned if the auditor of the future will issue a complete assurance opinion for the audited financial statements. The data analytics working group addressed this issue in their request for proposals and came to the opinion that the auditor is not able to provide more than "reasonable assurance" (IAASB, 2016).

Another open question for some practitioners was what the audit documentation needs to look like when applying data analytics. For all audit procedures, according to ISA's documentation requirements, an experienced auditor must understand the steps performed and the conclusion drawn when reading the documentation for each audit procedure. According to Küng et al. (2017), using data analytics requires the auditor to document information about the data used, proof of complete, correct and consistent data transfer, description of applied analysis such as filtering and aggregating, finalized with the interpretation of the results, and the respective conclusion.

5.3 Summary

Based on the interviews conducted, the impression was gained that data analytics is currently used to supplement an audit. Moreover, the use of advanced data analytics tools in the audit profession is in its infancy. When the collected applications are mapped onto the audit process, the following picture emerges (Figure 18).

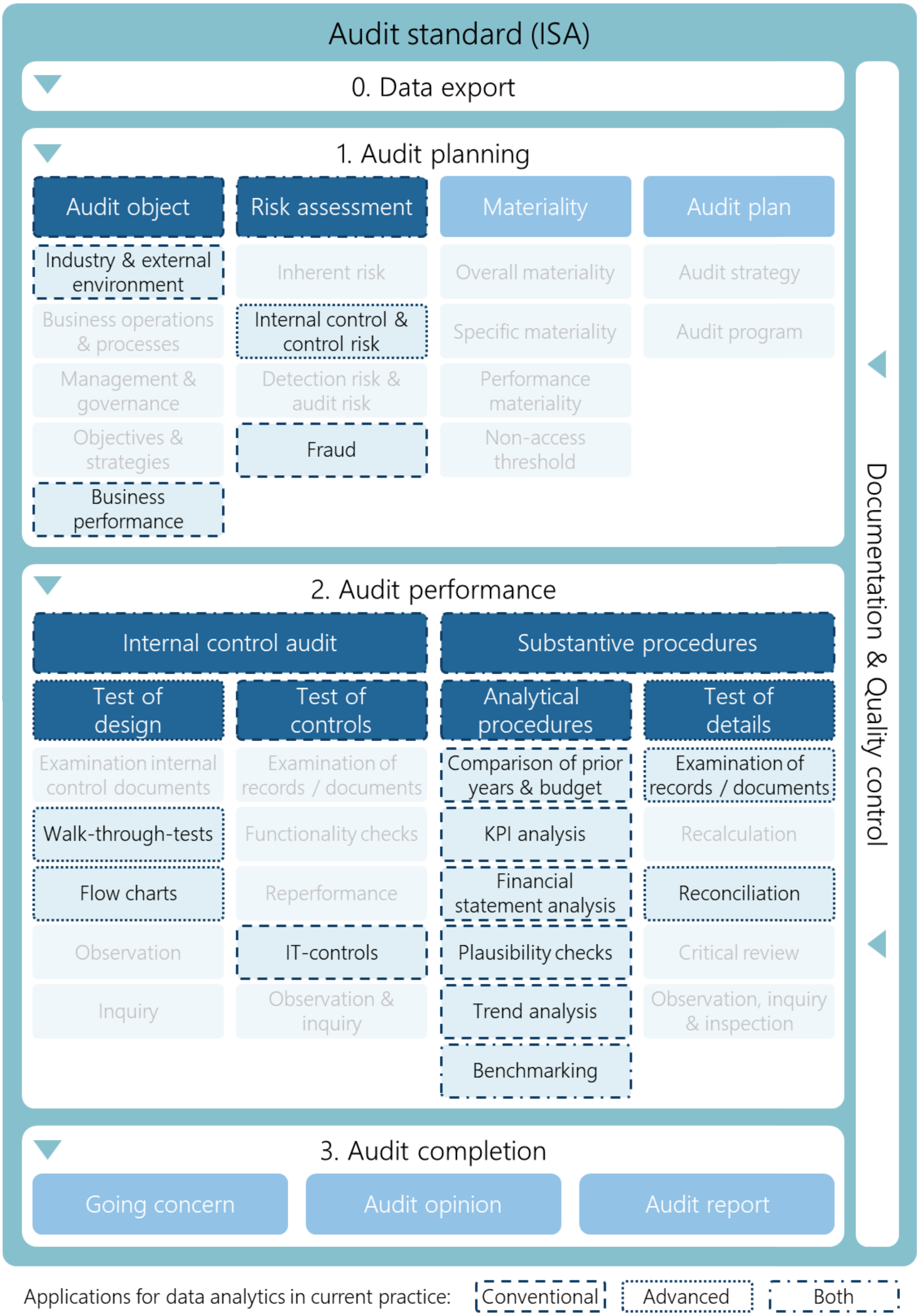

Figure 18 Illustration identified applications in the current audit practice and changes

6 Future applications

6.1 Research methodology

This chapter addresses the third research question:

What are potential future applications for data analytics methods in external auditing?

For answering this research question, experts are the primary source of information.

As discussed in the previous chapter on current applications, conventional data analytics methods – including query-based tools to answer specific audit questions – are already widely used within the interviewed audit companies. However, advanced data analytics methods are only rarely used. According to Westermann and Spindler (2017), machine learning is barely a topic in auditing practice, a finding backed in the interviews conducted with the auditors in the last chapter.

Therefore, this chapter focuses on showcasing possible use cases of machine learning methods based on expert interviews. The inclusion criteria for an application is based on the Gartner hype cycle (Figure 19). Each application should have not yet reached the “plateau of productivity” where mainstream adoption starts to take off, and where the use case’s “broad market applicability and relevance are paying off” (Gartner, 2019).

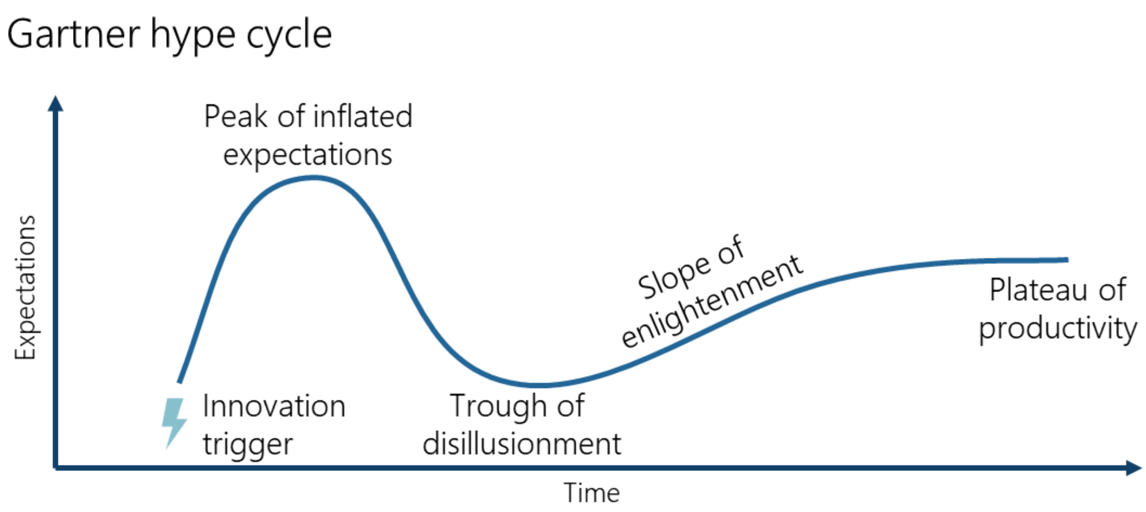

Figure 19 Gartner hype cycle retrieved from Gartner (2019)

The selection of the applications subsequently profiled is not comprehensive but is intended to exemplify and illustrate the potential that advanced data analytics technologies bring to the audit industry for every audit company performing a full statutory audit.

6.1.1 Sample selection

Bogner and Menz (2009, p. 54) define an expert as a person that "has technical, process and interpretative knowledge that refers to a specific field of action, by virtue of the fact that the expert acts in a relevant way". Guidelines to measure the quality of each expert do not exist at the moment, and establishing such criteria can be a complicated endeavor (Gläser & Laudel, 2009). Meuser and Nagel (2009) state that each researcher must decide on the criterion for calling someone an expert in a particular field. For this dissertation, an expert is defined as a person who has knowledge about advanced data analytics methods and is confronted with this topic daily. Ideally, the expert also has auditing experience or knowledge. In cases where the latter criterion was not fulfilled, auditing and the auditing process were briefly explained to each expert before the interview started.

Given that this is a highly specific target group, the “snowball sampling” method was used to get access to more experts. Brewerton and Millward (2001) state that in this method, “initial participants in the research are asked to provide further contacts”. The authors recommend this approach when it is difficult to get access to a small and closed group. Nevertheless, this method introduces a potential selection bias, since individuals might name acquaintances, “who may share similar views” (Wilson, 2009, p. 199). For this research question, given its exploratory nature, this was a lesser concern.

Experts were identified by asking the auditors interviewed in chapter 5 (the snowball sampling method), through recommendations from the dissertation committee, and by contacting companies that offer data science consulting services or startups that build machine learning products.

The final selection of thirteen experts includes three data scientists with Big 4 experience, three data scientists currently working in an advisory capacity, two heads of digital transformation and AI, three providers of technology-driven audit services, one PhD student in computer sciences with a focus on AI, and one PhD student in capital market research who is also the founder of a startup focused on applying AI in business practice.

6.1.2 Data collection

The recommendation of Meuser and Nagel (2009) to conduct open interviews based on a topic guide was followed. As both authors state, “the open interview provides the room for the interviewee to unfold his own outlooks and reflections” (Meuser & Nagel, 2009, p. 31). The topic guide helps to avoid creating the impression of an incompetent interviewer and getting lost in topics that are not relevant to the research question (Meuser & Nagel, 2002).

Additionally, it is crucial for the interviewer to be recognized as a “co-expert” by the expert, which increases the chances that useful facts and information is shared (Bogner & Menz, 2009; Rubin & Rubin, 2012). To ensure this, the expert interviews were carried out after the literature review and interviews with auditors.

While most interviews were conducted in person, three had to be conducted by telephone. With the experts' permission, most interviews were audiotaped. Responses were anonymized unless permission was granted to identify the interviewee. Out of the thirteen interviews, two experts declined to talk about specific use cases to protect their developments within this field. However, they agreed to talk about implementation challenges, and their input was used to evaluate other suggested use cases.

6.1.3 Data analysis

The data accumulated through the expert interviews was analyzed according to the principles of Meuser and Nagel (1991). Under certain circumstances, they regard it as unnecessary to transcribe the audio recordings of expert interviews in full length, for example, when some elements are not important key factors to answer the research questions at hand, which is given for this research question. Following their recommendation, thematic transcripts (Alvesson, 2011) were created, focused on content about building use cases. By condensing the spoken content, the amount of data and its complexity is reduced to what is essential for the research question. After finishing each transcript, the audiotapes were re-checked to make sure no essential findings were lost.

Once the completeness of the review was confirmed, the content of the summaries was split into segments. By segmenting the content, it was possible to compare and evaluate expert opinions from different interviews.

With "cross-checking" (Dean & Whyte, 1979), it was possible to guarantee a critical reflection of the spoken content and create awareness where the use cases might not be feasible at the moment. By applying the cross-checking method, some applications proposed by the experts were eliminated as the suggested use cases were not feasible, not practical for the audit profession, or already implemented.

A particular focus was given to the practicability of the suggested use cases for today's auditor based on actually available datasets. Since the auditor has to perform journal entry testing (ISA 240), the journal entries are a first dataset the auditor has undeniable access to and therefore are a natural target for applying data analytics.

Out of the 13 interviews, five future use cases were identified. The source of two applications is directly stated in the text. The experts involved with those applications reviewed the written text. For the remaining applications, anonymity was requested. As each main idea was derived from one single interview, citations by anonymized abbreviations were removed to benefit readability, and the term "expert" was used instead.

Each use case is presented in the following general order: field of application, dataset, methods, and implications.

6.2 Journal entry testing with clustering

Field of application

This first application focuses on the auditor's duty to perform journal entry testing (ISA 240: Auditor's responsibilities relating to fraud in an audit of financial statements, and ISA 315: Identifying and assessing the risks of material misstatement through understanding the entity and its environment). The use case is situated within the fraud substage of the audit planning stage. The purpose of this analysis is to identify material misstatements and fraud by looking at outliers.

In the field of outlier detection, the differentiation between global and local outliers is made. Global outliers show unusual or rare journal entry characteristics such as timestamps during unusual business hours (Schreyer et al., 2018). Their unusual behavior makes global outliers easier to detect, whereas local anomalies are in danger of being hidden in mainstream bookkeeping flows.

As the expert mentioned, "intelligent" fraud tries to be as close as possible to a usual transaction to not trigger any alerts. Employees committing fraud often know where the internal control loopholes are, what kind of reports are created, and the validation procedures in regards to approval thresholds. These entries, which try to mimic the "usual" journal entry behavior as closely as possible, are also known as local outliers.

In this context, the expert suggested using clustering (unsupervised machine learning method) to identify outliers. Using this method has the advantage of avoiding to label the data with the target value, which would, in this case, be anomaly and non-anomaly.

Dataset

The dataset consists of all journal entries from the audit client within the analyzed audit period. The basic structure of journal entries is homogenous across companies. They include the debit and credit account, amount, text, value-added tax percentage, and date of the entry. ERP-systems additionally capture information such as type of entry (e.g., manual, automatic), posting key (e.g., 01 = accounts payable payments), document type, business unit, profit-center, and more, depending on the client’s customization of the system.

However, based on the expert's opinion, if bookkeeping is not fully automated and is performed by more than one person, the content can be different. For larger companies, accounting manuals exist to ensure accounts are used the same way in every company unit. Nonetheless, it is not guaranteed that these manuals are followed consistently. More significant variances can be expected in what is inserted in the text field of a journal entry. Such inconsistencies create noise within the dataset and can complicate the interpretation of the data itself and the results. However, according to another interviewed expert, this also provides a first red flag for the auditor, as he would see that accounting is performed inconsistently, despite the presence of an accounting manual.

To reduce the number of variables to the most useful ones, the expert suggested using principal component analysis. This analysis helps to answer the question of which attributes explain a significant amount of variation contained in the present dataset, by performing a dimensionality reduction of the dataset (Kotu & Deshpande, 2014). To select the relevant attributes, a variance threshold (e.g., 95%) is defined by the modeler. Attributes are added one by one, in order of incremental explanatory power, until this threshold is reached.

The expert suggested using a maximum of four attributes for clustering, because fewer attributes make it easier to interpret the results. According to Abbott (2014), principal component analysis works best with numeric data. This condition should be fulfilled by the journal entries, except for the text field. Where non-numeric variables are present, a solution could be to either recode nominal variables or to calculate the distance with another algorithm (Torgo, 2011). Principal component analysis is highly sensitive to scaling effects, which can be minimized by first normalizing the numeric input data (Han et al., 2012; Kotu & Deshpande, 2014).

Therefore, already when preprocessing the dataset, care must be taken to tailor the method to the individual dataset characteristics. Even in this early stage, interesting insights could appear for the auditor.

Methods

Once the major variables have been identified, the model can be created. Unfortunately, the expert did not mention a specific clustering method, and a variety of clustering algorithms exist. In this absence, the literature was checked for potential methods.

In general, Kantardzic (2011) recommends testing the given dataset first for clustering tendencies, as every clustering algorithm will produce clusters regardless of whether they are present or not. Once this is performed, two clustering algorithms for outlier detection suggested by Torgo (2011) may be promising approaches:

Local Outlier Factors (LOF)

Breunig, Kriegel, Ng, and Sander (2000) developed a local outlier factor (LOF) method, which is considered a state-of-the-art outlier ranking method (Torgo, 2011). The suggested clustering method focuses on the local anomaly type and is therefore well suited for this application. Based on Breunig et al. (2000), each journal entry is assigned a degree of being an outlier. This degree represents, for each instance, how isolated it is from the surrounding neighbor instances. The method makes use of the local density of the instances, where instances in low-density regions are defined as outliers (Torgo, 2011). From the calculation of the local outlier factor, a ranked list can be created to give the auditor guidance on which outliers to check first.

Breunig et al. (2000) successfully tested their method with 64 attributes. Since the method can handle so many attributes, the question arises whether the principal component analysis needs to be applied. However, the principle still applies that the more attributes are used, the more difficult it becomes to describe or explain why an identified instance is an outlier. This especially concerns high-dimensional datasets, as local outliers "may be outlying only on some, but not on all, dimensions" (Breunig et al., 2000, p. 103; Knorr & Ng, 1999). This can make it challenging for the auditor to select journal entries for review and argue why they were chosen. However, if fraud was uncovered with such a method, another expert suggested that "rule-based" justifications could be derived subsequently.

Torgo (2011) used this method as well. He showed that the modeler needs to specify how many neighbors are used to calculate the local outlier factor for each instance and to determine a density threshold.

Clustering-based outlier rankings

This method is based on a hierarchical agglomerative clustering algorithm (Torgo, 2007) and does not differentiate between local and global outliers. This type of clustering algorithm assumes that each data point represents its own cluster (Roiger, 2017). In each iteration, the algorithm merges the two most similar clusters with the final iteration merging all instances into one cluster. Which similarity measure is applied depends on the selection of the modeler. Torgo (2007) selected Ward's agglomeration method (Ward, 1963), which merges groups based on maximum similarity (minimum variance).

As explained before, outliers are presumed to differ from other transactions. Consequently, outliers do not show a high similarity with others and therefore are selected at a later stage of the iterative merging process. The iterative process is depicted in a dendrogram, which shows which clusters have been merged during which iteration. The assumption of outliers merging at a later stage leads to a model flaw. In the case of multiple outliers, they are merged as a group and, therefore, may not be detected by the suggested model (Torgo, 2007). To address this issue, Torgo (2007) added an outlier score to the model, which considers the fact that a small outlier group will probably be merged with a much larger cluster towards the end of the clustering process. Based on that score, a ranking can be created.

In a different application, Westermann and Spindler (2017) discovered that no single clustering algorithm was able to identify all anomalies in their artificial test set. However, by combining ten clustering algorithms and introducing the ensemble voting method, the results improved significantly. For choosing the right distance measure for the presented method, a similar approach could be used. Applying this approach to the presented method, various similarity measures could be used, with the ensemble voting method giving the final decision. This would yield the cluster to which the data point (journal entry) belongs, or which clusters are merged.

Implications

Both of the methods detailed above allow the derivation of an outlier ranking, which in practice then turns into a prioritized list of transactions for the auditor to review. This is helpful for the auditor by giving guidance on where to start.

An advantage of clustering is that no significant amount of data is needed. Therefore, it can also be used for smaller audit engagements. However, which of the presented methods is effective, or if even a combination of both needs to be applied, still needs to be tested in practice.

The model itself needs to be built by a data scientist because many decisions are required in the process. Once the results arrive, either further clustering is necessary (e.g., if there are too many outliers) or a manual follow-up on the identified journal entries needs to be performed. Considering that this method has the potential to unearth a significant number of anomalies, the amount of effort required by the auditor during the audit could be significantly increased. In this case, the auditor would be responsible for determining a reasonable threshold by applying his professional judgment.

Additionally, it should be noted that the application of this method leads to a review of a "potential" anomaly. Up until the detailed review, nothing is proven. There may still be a reasonable business justification for the discrepancy.

6.3 Journal entry testing with adversarial autoencoder neural networks

Field of application

The following application also focuses on journal entry testing, which concerns the auditing standards ISA 240 and ISA 315 as mentioned before. Therefore, this method is part of the audit procedure "risk assessment" in the audit planning stage. The goal is to detect anomalies in the form of fraud and material misstatements.

As observed, finding fraud is a challenging endeavor, primarily since existing fraud cases are vastly outnumbered by nonfraud cases in any dataset. One approach that surfaced during expert interviews is the use of an adversarial autoencoder network combined with an anomaly score to detect deviations within journal entries.

The presented application has been tested with real-world journal entries. The first use case was published with the title "Detection of anomalies in large-scale accounting data using deep autoencoder networks" by Schreyer, Sattarov, Borth, Dengel, and Reimer (2017). The application was extended with an adversarial component in a follow-up paper with the name "Detection of accounting anomalies in the latent space using adversarial autoencoder neural networks" by Schreyer, Sattarov, Schulze, Reimer, and Borth (2019b). Both methods are from the deep neural network algorithm family (unsupervised learning) (LeCun, Bengio, & Hinton, 2015).

Especially for audit engagements concerning larger corporations, auditors are confronted with millions of journal entries, which need to be tested for fraudulent actions or material misstatements with limited personnel, resources, and time (Schreyer et al., 2018). Although rule-based tests have been implemented in the audit process, they are limited to known-fraud scenarios and rely on the creativity of the auditor in defining the rules. Using adversarial autoencoder neural networks promises to be able to identify previously unknown fraudulent patterns.

Typically, neural network methods are black-box by nature, meaning that the network does not provide a humanly comprehensible reasoning as to why a specific transaction is seen as an anomaly. Based on this issue, the expert suggested using an adversarial autoencoder neural network combined with an anomaly score to overcome this issue. The chosen architectural structure tries to learn a "semantic meaningful representation of the journal entries" (Schreyer et al., 2019b, p. 2), which allows the trained network to partition the journal entries into meaningful groups based on shared characteristics.

Dataset

The required dataset consists of all the journal entries of an individual company. As adversarial autoencoder neural networks are generally "data-hungry" methods, a massive number of journal entries is needed. While no specific minimum number can be named, the presented application used over 300'000 journal entries (Schreyer et al., 2019b). As with any other data analytics method, the caveat applies that the trained model is only as good as the quality of the underlying dataset. Since fraudulent journal entries occur rarely compared to non-fraudulent journal entries, the challenge is to train the network in a way that differentiates anomalies from noise within the presented dataset (Schreyer et al., 2018).

For the proposed method, all journal entries are transformed into a numerical feature space to be processable for the model. This procedure is called one-hot encoding. Categorical attributes (e.g., document type) are transformed into a column, and a "1" is entered if that document type was used for the journal entry and "0" if not. Figure 20 shows an illustrative example of this process.

Document type
Sales invoice
Purchase invoice
Outgoing payment

Sales invoice	Purchase invoice	Outgoing payment
1	0	0
0	1	0
0	0	1

Figure 20 Example of one-hot encoding based on Knox (2018)

To the dataset, local and global anomalies are manually added, and each journal entry is labeled accordingly (synthetic global anomaly, synthetic local anomaly, or non-synthetic regular entry), but only for visualization at a later point. The anomalies were entered in proportion to the occurrences in a real-world setting (lower than 0.02%).

Method

Overall, the method comprises three jointly used, distinct neural networks. Adversarial autoencoders (Makhzani, Shlens, Jaitly, Goodfellow, & Frey, 2015) are inspired by the same architecture as deep neural networks, which have been explained in chapter 3.4.3. Adversarial autoencoders have the advantage that no labeled data is needed; as such, they are unsupervised learning methods.

The architectural structure of the adversarial autoencoder neural network is shown in Figure 21[11].

[11] For further details, please refer to the publication of Schreyer et al. (2019b).

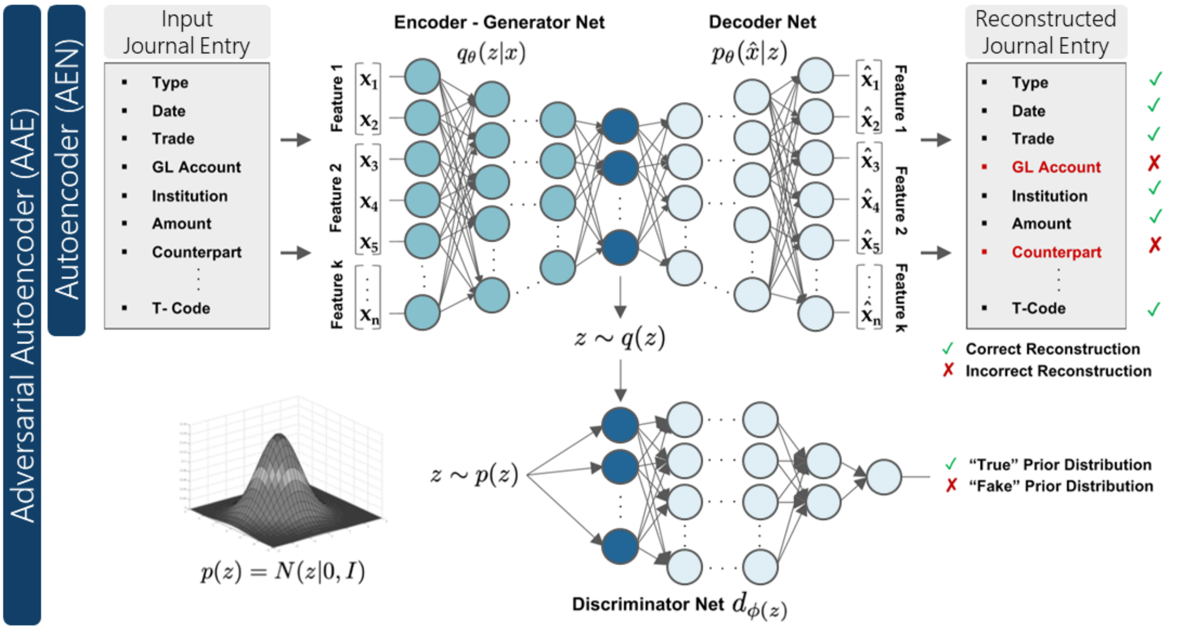

Figure 21 Adversarial autoencoder architecture retrieved from Schreyer et al. (2019b, p. 3)

Autoencoder neural networks are also known as replicator neural networks because they contain an encoder net that compresses the input information into a latent space and from there the decoder net reconstructs the decompressed information into an output (Schreyer et al., 2017). Within the latent space, the decompressed information is forced through a bottleneck. The bottleneck prevents the autoencoder from just remembering all shown journal entries and instead forces it to learn the semantics of the dataset, which is needed to form meaningful clusters (groups). To make sure the semantics are learned, the expert makes use of the distributions and the dependencies between the journal entry attributes (e.g., document type, date, amount, general ledger accounts). For example, the posting key "outgoing payment" will occur many times together with a general ledger bank account. As such, these attributes possess a strong dependency. An optional enhancement in this phase is the adversarial part (discriminator net – third network), which improves the performance of the autoencoder.

Once the decoder net has reconstructed the compressed information, the original input can be compared with the output. The difference between the reconstructed output and the original input is called "reconstruction error". During the autoencoder training phase, the model aims to minimize the reconstruction error. The model creator determines the length of the training phase by setting the number of epochs. The more epochs of training are carried out, the more time the network has to find the optimal minimum for the reconstruction error.

Additionally, an anomaly score is created to establish a ranking of the journal entries (Schreyer et al., 2019b). The anomaly score combines the mode divergence error and the reconstruction error. Mode divergence error measures how far away each journal entry data point is from the mode of the respective cluster, assuming that the further away a journal entry is from the respective mode, the more suspicious it is. As mentioned before, the reconstruction error is the difference between the reconstructed output and the original input.

Besides determining the number of training epochs, the α (alpha) parameter (determining the weight of the reconstruction error in the anomaly score) and the number of semantic clusters must be determined as input factors. Choosing adequate parameters is an exhaustive and systematic exercise that relies on the experience of the data scientist. When choosing five clusters, α of 0.8 and 5'000 training epochs, the following results appeared for tested dataset:

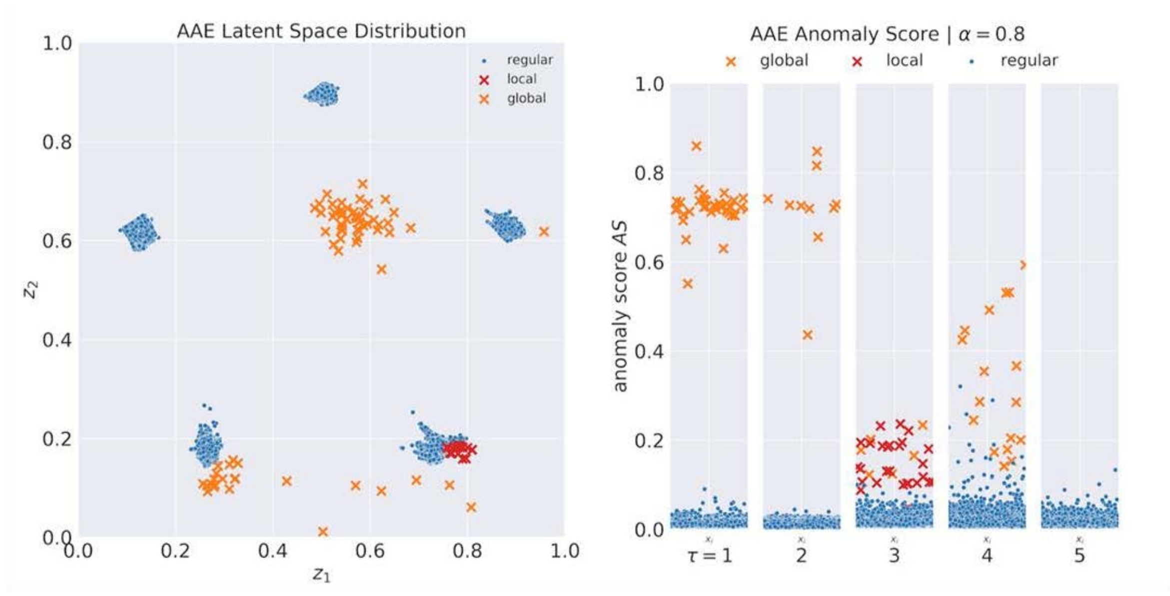

Figure 22 Latent space representation of trained adversarial autoencoder neural network retrieved from Schreyer et al. (2019b, p. 11)

The results of the model show that it is capable of learning the semantic partitions of a given journal entry dataset, as represented in the visualization on the left-hand side of Figure 22. Each cluster combines journal entries with similar characteristics (e.g., customer invoices, payments, manual entries). The right-hand side visualization shows every single cluster and its anomaly score. Each data point within the cluster represents a journal entry and its corresponding anomaly score. As the results indicate, the network identified both the local and global anomalies.

Implications

In practice, the data scientist conducts such an analysis, and the auditor receives the results as depicted in Figure 22. In this visualization, all data points would be shown in the same color as the auditor does not yet know the anomalies within the dataset.

According to ISA 530.4, the main objective of sampling is "to provide a reasonable basis for the auditor to draw conclusions about the population from which the sample is selected". Based on this ISA regulation, the interviewed expert suggested that the auditor starts checking the features of each cluster's centroid (middle point), which is the representative of all journal entries within that cluster. By looking at each centroid, the auditor gets a basic understanding of what the cluster represents (e.g., customer invoices) and its characteristics (e.g., document type RE, posting key 31, transaction code 04). Based on the auditor's judgment on this first assessment, further investigations may be necessary.

A point for discussion is the acceptable distance from the centroid, before considering a journal entry an anomaly. The expert recommended the following mathematical rule of thumb: outliers are two to three standard deviations away from the mean or in this case, the centroid. Every journal entry outside this range can be regarded as a global anomaly.

However, this would not uncover local anomalies. The expert assumes that local anomalies are connected with the reconstruction error, meaning that the adversarial autoencoder neural network has difficulties with reconstructing the original input (journal entry). The auditor can retrieve a ranked list with the highest reconstruction errors, or increase α to a high level to give more weight to the reconstruction error within the formula of the anomaly score. The anomaly score also includes the distance from the centroid, which applies more to global anomalies, and therefore the anomaly score needs to be recalibrated when focusing on local anomalies.

Depending on the number and distribution of the global and local anomalies found, further clustering can be conducted. Another possibility is adding the transaction value of each identified anomaly. Both methods might be worthwhile in order to work faster through the identified anomalies.

When applying such a method and conducting certain follow-ups, the question must be asked whether the audit procedures performed provide the necessary reasonable assurance for the whole population.

The auditor is required to prepare a sufficient and appropriate documentation (ISA 230). Especially if audit evidence is generated from such a method, another experienced auditor needs to be able to understand how this evidence was obtained (ISA 230.8) and any conclusions reached (ISA 230.13b). This is a major hurdle to the application of this method. It is already demanding to build such a model, but to document everything presents a further challenge, especially with the condition that any experienced auditor should be able to understand it.

A further limitation of the proposed method is that it can only be used for larger audit engagements where thousands of journal entries are available, as adversarial autoencoder neural networks are data-hungry models. Building such a model is expensive, as special hardware (e.g., with high processing power) is required to run the models. As an alternative to local hardware, cloud-solutions are possible, but this requires the audit client's permission to upload the data to a cloud.

Unfortunately, such an application can also be abused to generate real-looking fake journal entries, that can be posted to the ERP-system, as shown by Schreyer, Sattarov, Reimer, and Borth (2019a). The same model can therefore be used not only to uncover fraud, but to also commit fraud by constructing journal entries which are even more challenging to detect. In their paper, the authors also provided evidence that current state-of-the-art Computer Assisted Audit Techniques (CAATs) used in external auditing have difficulties detecting the generated fake journal entries.

6.4 Journal entry testing with association rules

Field of application

The following approach again focuses on journal entry testing. Therefore, the following use case falls into the audit planning stage within the fraud substage.

This application represents an enhancement of an existing rule-based method. The expert called the existing method a "knowledge account application", which is used to identify unusual journal entries. The chart of accounts from the audit client is transferred to a knowledge database and mapped to an existing chart of accounts template, considering only the accounts used during the fiscal year under consideration. Each industry has some accounting specifics which should be considered (e.g., pharma – research and development; airline – leasing). Based on the client's industry, assumed business-usual and business-unusual journal entries are stored in the knowledge database. For example, a booking of trade receivables to sales is a usual journal entry, whereas a booking of bank liabilities to sales should trigger a red flag.

Before running the test, these predefined assumptions of business-usual and unusual journal entries are cross-checked with the client, as some relationships might appear unusual to the auditor but happen in the daily business of the client for a legitimate reason (Earley, 2015). Once the veracity of the assumptions has been confirmed with the client, the auditor can run the test on all journal entries and uncover which unusual journal entries occurred during the fiscal year.

As this method is limited to the set rules of the auditor, the expert suggested an explorative pattern recognition method. The expert recommended that the model be created by a data scientist, as the results can prove deceptive to a novice (Larose & Larose, 2015).

Dataset

The dataset consists of all journal entries within the fiscal year. The auditor or data scientist should check if the journal entries handed over can be reconciled to the financial statements presented (ensuring completeness).

The expert considered the timestamp and transaction amount as the two most useful attributes because the algorithm is sensitive to the flow of a transaction (e.g., receiving order, shipping, sending invoice). This assumes that the pattern in the journal entry data does not change abruptly unless something abnormal happens (Aggarwal, 2017). With this assumption in mind, it follows that the journal entries of companies within a group cannot be mixed, because otherwise the time-series data would be diluted. Performing the analysis for each company could also reduce the inherent bias of the data since each accountant might interpret the chart of accounts slightly differently in practice. Nevertheless, if the temporal continuity of the dataset is not given, no algorithm with this prerequisite can be used.

As a more accessible alternative for auditors, the expert suggested using an association rules method based on the account combinations used within the journal entries of the current year.

Method

As the association rules method was explained in chapter 3.4.3.2.1, only additional knowledge will be discussed in this subchapter. The association rules analysis could be carried out with the apriori algorithm.

Depending on the structure of the audit client's chart of accounts, it might be useful to combine certain accounts. For example, if an audit client has a subaccount for each customer and supplier, it is worthwhile to combine these accounts into a generic trade receivables or trade payables account. For the proposed implementation below, only the debit and credit position of each journal entry will be given to the algorithm. Alternatively, the chart of accounts could also be mapped to a template as it is done in the rule-based method previously described.

An illustrative example of what the output of the algorithm could look like is shown in Table 8.

Debit position	**Credit position**	**Support**	**Confidence**
Trade receivables	Sales	0.3	0.90
Salary expenses	Liabilities to employees	0.1	0.95
…	…	…	…
Bank liabilities	Sales	0.0000008	0.000005
Bank liabilities	Equity	0.0000006	0.000003

Table 8 Illustrative example of association analysis results

"Support" indicates how many journal entries this rule applies to in the given dataset (Aggarwal & Han, 2014). The journal entry in the first row, "trade receivables to sales," covers 30% of the whole journal entry population. "Confidence" defines (Kantardzic, 2011) how probable it is, when a specific "debit position" occurs, that a particular "credit position" follows. The confidence value in the first row means that in 90% of the analyzed journal entries where trade receivables are on the debit side, sales are, in turn, on the credit side. The next journal entry in row two also seems explainable to the auditor. However, at the end of the list, two unusual journal entries can be found. Those should be checked with the receipts/documentation behind these entries and eventually discussed with the audit client. "Bank liabilities to sales" could be an erroneous journal entry. "Bank liabilities to equity" can be found in practice, but it is rare. It could mean the bank changed its debt into equity, which is generally only done in times of financial distress. This would be a red flag for the going concern assumption and has a material impact on the presentation of the financial statements.

Table 8 is just a small excerpt from a results table. In practice, a longer list will emerge. The length of the list depends on the number of unique journal entries within the dataset. As the expert mentioned, retrieving the results is the easy part; the hard work starts when association rules turn out not to be conclusive. Lantz (2015) confirms that some rules are inexplicable, making it nearly impossible to figure out how the retrieved information should be used. Additionally, the auditor generally does not have unlimited time to follow-up on each discovered pattern. In such cases, the association rules can be segregated further into actionable and non-actionable clusters.

Another possibility to reduce the patterns identified in the literature (e.g., Kantardzic, 2011) is to set a threshold for support or confidence. Usually, both are set to a higher level to find the most frequent rules, but an auditor would want to have a look at the most infrequent patterns, and so eliminate the ones with the highest support or confidence.

However, the recommendation from the expert is not to limit the results in order to have a full transparent perspective, but rather to filter the identified association rules with other methods. A useful association rules reduction method for the auditor could be combining the results of this analysis with the existing rule-based knowledge account application, to eliminate the business-usual patterns. Another reduction opportunity is to create a weighting score by summing up all transaction amounts for the identified patterns. This would help to focus only on material unusual journal entries.

Many permutations of the suggested method can be applied to a dataset. For example, more features (e.g., journal entry type) could be included in the analysis, or the results of alternative association rule algorithms could be tested.

Implication

As the test is performed using the entire population of journal entries, a transparent view is gained in the process. Thoroughly scanning the journal entries with the presented method helps the auditor to reduce the risk of material misstatements or fraud in the financial statements by uncovering unusual account combinations. Additionally, with this method, the whole population of journal entries is reviewed, which gives the auditor more comfort in terms of audit assurance. However, this only applies if the auditor first reconciled the journal entries provided with the financial statement information to make sure completeness is given.

A potential drawback of the method is that the expert doubted if any efficiency gains can be realized, as the follow-ups can be challenging, time consuming and demanding. Furthermore, there is a tendency that the more journal entries are provided, the longer the list of rules becomes. Beyond a certain quantity of journal entries, additional methods are needed to guide the focus of the auditor to the important results.

An advantage of the method, on the other hand, is that it can be applied to any audit client, as it works with any number of journal entries. Also, for a basic analysis of the journal entries, it can be assumed that the necessary data quality is present, as the journal entries are also the baseline used for creating the financial statements. Therefore, this approach is less susceptible to data quality issues than other methods.

A crucial point for the success of this method is that when several sub-accounts are manually combined into one (e.g., replacing all sub-receivables accounts with the main receivables account), no errors are incorporated, which would lead to unnecessary follow-ups.

6.5 Forming predictions for analytical procedures

Field of application

The following use case falls into no particular audit stage as analytical procedures are used in every audit stage. However, it is probably more applicable for analytical procedures in the audit planning and audit performance stage. Analytical procedures are applied to assess the consistency and economic plausibility of expected values (Marten et al., 2015).

Before performing analytical procedures, the auditor needs to set an expected value as well as a deviation tolerance range (Marten et al., 2015). For setting the expected value, the auditor can use the support of predictive machine learning algorithms. The basic idea is to predict an expected value and test it against the observed value. Differences between the values provide a basis for further investigation into why this is the case.

Three auditors interviewed in chapter 5 named predictions as a data analytics method likely to be valuable to the auditor. Within the data analytics literature, predictions can have two meanings. Within this subchapter, the definition of prediction from Roiger (2017, p. 36) is used, saying that the "purpose of a predictive model is to determine future outcome rather than current behavior". However, predicting an outcome in terms of assigning an instance to a class (e.g., fraud vs. nonfraud) is also a form of prediction (Kelleher & Tierney, 2018).

Using such prediction methods can help evaluate the business performance during the audit planning phase, and is especially useful for trend analysis, plausibility checks, benchmarking, and financial statement analysis during the substantive analytical procedures in the audit performance stage.

Dataset

In the expert's view, about 80% of the whole time required for creating the model will be occupied with the data preprocessing stage. Regarding the dataset, the expert suggested using both internal (e.g., financial data from previous years) as well as external data. For data preprocessing, the expert recommended looking at the correlation of numbers and accounts, sales in terms of company size and number of journal entries.

Regarding the expert's suggested use of external data, various commercial offerings by data providers exist. Based on the interviews with auditors, some audit companies already collect data about the audited company's industry and competitors to gain a better understanding of the ecosystem of the audit client (see subchapter 5.2.3.1). Buying external data to validate the present results is already common practice for some audit companies.

Method

When questioned about adequate methods to create the predictions, the expert explained that many methods can be used. He suggested testing various methods such as random forests, linear/logistic regression, and neural networks with real numbers, to discover the most useful one for the auditor. For neural networks, it is important to consider that a vast amount of data is needed. If one method has a considerably better accuracy rate in terms of making predictions, then it should be used. From his experience, however, there is no significant difference in accuracy between the methods, and therefore, the simplest or more user-friendly one is preferable for implementation.

Another possibility which focuses more on the time aspect of the data is time-series algorithms such as ARIMA. ARIMA stands for "autoregressive integrated moving average" and is a popular method for times series modeling that is well established in the field of econometrics (Wu & Coggeshall, 2012). According to Kotu and Deshpande (2014, p. 317), "autoregression methods are basically regression models applied on lag series where each lag series is a new predictor used to fit the dependent variable”.

Such a method might be especially useful for related financial accounts, as a major feature of the method is that the next value is related to the most recent one (Wu & Coggeshall, 2012). Concerning those relationships, Vandervelde (2006) differentiates between cross-section and temporal interactions within accounting data. The former are relations such as sales and accounts receivables, which might be less useful for the method. The latter refers to the time-connectedness. For example, a wrongly depreciated asset will be carried from one period to the next until the asset is sold or completely depreciated. Due to the connectedness of the accounts, a past error in one account will be carried over to the next period.

Kotu and Deshpande (2014, p. 317) write that “ARIMA is a complex technique and it requires a great deal of experience to produce good forecast results”. This indicates that the data preprocessing phase and the available supporting techniques (e.g., box-cox transformation or smoothing) for this method should be considered (Wu & Coggeshall, 2012). Depending on the variance of the dataset, ARCH (autoregressive conditional heteroscedasticity) or GARCH (generalized autoregressive conditional heteroscedasticity) might be more appropriate.

It is recommended to feed the model the new company data once the audit is completed to continuously improve model performance and make sure the predictions are improved. In addition, Kotu and Deshpande (2014) suggest involving forecast errors as another attribute for the models. Another aspect to consider is how often re-training the model occurs, as some companies also have quarterly and half-year financial statements.

Independent of which algorithm is chosen, the end result will be an account value prediction or a prediction of a financial ratio. However, the reliability of the prediction remains subject to the auditor's professional judgment. Combining the predictions with a visual chart can assist the auditor's assessment of the expected value for performing analytical procedures.

An illustrative example for such a visualization is provided in Figure 23, based on the CaseWare AI tool. CaseWare, issuer of the IDEA audit software, provides a publicly available prediction tool. However, this tool covers only public companies from the United States. Additionally, it appears that this tool is at an early stage of development, since some features (e.g., services industry ratios) do not always work, as can be seen in Figure 23. CaseWare keeps the methods used to create the predictions proprietary and only reveals that a collection of machine learning techniques (CaseWare, n.d.) is used, which is a generic description.

In the example, four KPIs of the CaseWare AI tool for Microsoft are shown, depicting yearly values. As three quarterly reports of 2019 are already issued, the predicted values for the year can be compared with actual values to evaluate the reliability of the prediction itself. As the service industry ratios are not working consistently, the comparison is made solely of the existing ratios and their predictions.

Figure 23 Yearly financial ratios of Microsoft based on CaseWare (2019)

The first two examples (owners' equity ratio and debt ratio) show a positive example of predictions that are close to the observed values. A critique could be that in the past, there was never much variation of those two financial ratios. As such, making predictions is "easier". The lower two examples (accumulated depreciation to gross property plant and equipment as well as working capital) are shown as negative examples and predictions that should be critically questioned. Especially for working capital, there is a material difference between the prediction of around 71 billion US-Dollars and the actual value of around 111 billion US-Dollars.

Implications

The example illustrates both extremes: trustworthy and untrustworthy predictions. However, in practice, the most difficult cases are the ones in between. The auditor then has to decide if the prediction should be used for the analytical procedure or if own estimates should be derived. ISA refers to the auditor as the person in charge of making the estimates. If a prediction derived from an algorithm is unreasonable, the auditor is liable for trusting the prediction, as it is still the auditor's final decision. However, ISA 330.A10 states that analytical procedures alone do not provide sufficient appropriate evidence, which limits the burden on the auditor's shoulder in this case as further investigations are required anyway. To sum up, it can be concluded that prediction methods for analytical procedures are only supplementary tools for the auditor.

An advantage of the method is that such predictions can be performed with the historical data of a client. They also give the auditor some guidance and might be better than just using the budget of the client as expected values. Furthermore, for the first tests of this method, no new tools need to be acquired since off-the-shelf office software already provides add-ins such as XLMiner which can perform such predictions. This allows any audit company to experiment with the data at almost no additional cost. An additional advantage is that the numeric output is easily interpretable by an auditor.

On the downside, building more complicated models with larger data input amounts can be a demanding endeavor. Frequently, the auditor will need the help of a data scientist, as after each audit is completed new data is available to refine the models. Regarding the question of what kind of data (e.g., economic data, social media data) or features are useful for the models, research can help practitioners to find an answer.

In a broader sense, using prediction methods could also be helpful for risk assessment and going concern assumptions. For the latter, we have seen some tests already in chapter 4. For the risk assessment use case, the biggest hurdle would probably be choosing the appropriate attributes. In addition, either a method for non-numeric predictions is applied (e.g., low, medium, high), or a numeric scale is used. Larger audit companies especially could make use of a risk assessment prediction method, as such predictions could be based on their own data from the past. Therefore, such a project could be carried out independently from the audit client.

As shown, predictions can be used not only to help the auditor arrive at expected values for analytical procedures, but also in other areas within the audit process.

6.6 Audit workforce planning with machine learning

Field of application

The next use case is not used during an audit engagement itself but influences the audit process and is an example how machine learning can be used within the audit company more broadly. It is a tool that supports the auditor during audit planning.

The auditor is obliged by ISA 300 to plan the audit engagement, which explicitly includes "the nature, timing and extent of direction and supervision of engagement team members and the review of their work" (ISA 300.11). To make sure this requirement is fulfilled, the auditor needs to plan adequate staff assignments for the audit engagement. "Adequate", in this context, means scheduling the right people with the required hours at the right time.

For example, an experienced auditor rarely sends out the bank confirmation letters, as this is usually performed by support staff. At the same time, an inexperienced auditor would not be assigned to discuss provisions or goodwill with the client, as this is a task that requires more training and experience.

Since audit companies will have several ongoing audit engagements in parallel, performing the planning can be tricky, especially since leave, vacations, mandate preferences and staff experience levels need to be taken into consideration. In practice, spreadsheets, planning software, or a mix of both are currently used to perform workforce planning at audit firms, according to the expert.

Solving this complex multivariate optimization problem is a task that is both challenging and labor-intensive. The proposal is to apply machine learning to this task. A Swiss startup named Aspaara Algorithmic Solutions AG has developed this idea into an audit industry-ready solution and tested its usability with a Big 4 audit company.

Dataset

At least one full year of planning data is needed for training the algorithms, but ideally up to three years of data should be available. Based on the initial experiences of using the tool in practice, data quality is critical to ensure the quality of the outcome. If an appropriate level of data quality is not available, manual adjustments can be made to improve the data quality (e.g., correcting typos, standardizing spelling). No specific data format requirements are set, as the data is available either as a spreadsheet or as an export from previously used planning software.

Method

The operational sequence of this approach is depicted in Figure 24. The approach uses a data-driven optimization for rostering.

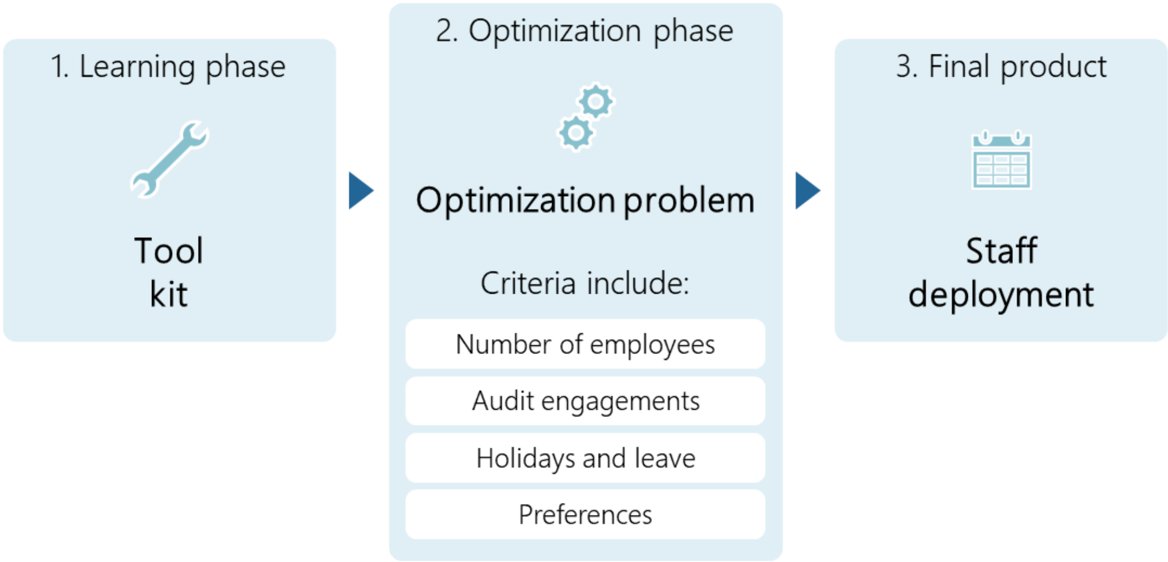

Figure 24 Operational sequence of workforce planning (own illustration)

The approach consists of two phases, namely the learning phase and the optimization phase. In the first stage (learning phase), a tool kit with various machine learning approaches is used. The appropriate machine learning method is chosen depending on the requirements and needs of the individual audit company.

In the learning phase, learning the staff rostering from the past is the primary goal. This is achieved with methods such as clustering to recognize patterns, or classification to learn about the appropriate allocation. Another option is to use algorithms for time series prediction. Which method is used, or if even a combination is ideal, depends on the needs and existing data of the individual audit firm. Once the algorithms have absorbed the specifics of past workforce planning, the learned insights are passed on to the next step, the "optimization phase."

For the optimization phase, the fixed optimization criteria of the upcoming audit season are inserted into the model. Solving the scheduling problem requires the following information (Sauer & Schumann, 2007):

- Required personnel per audit engagement in hours, and required level of experience
- Available staff in terms of hours and expert knowledge
- Hard constraints (e.g., maximum working hours, minimizing travel time)
- Soft constraints (e.g., employee preferences for specific clients or industries)
- Goals such as maximum acceptable overtime, target workforce utilization

Hard constraints are handled as a priority when solving the problem, whereas soft constraints are considered if possible. Once the data has been fed into the system, the model solves a global optimization problem. The final output is a staff deployment plan for the upcoming audit season. Depending on the chosen set-up of the audit firm, the results can be depicted as a dashboard combined with KPIs with a drill-down option or as a spreadsheet.

Further details could not be provided to safeguard the intellectual property of the founders.

Implications

The expert claimed that several benefits were realized when the system was implemented. KPIs like capacity utilization and travel time have improved, leading to cost savings for the audit company. Including the preferences of the workforce improved employee satisfaction. The planning process takes less time as the system performs the work, making the process more efficient and less prone to error. No additional system needs to be monitored as the tool can be incorporated into the existing IT environment. By implementing this tool, more useful output can be generated with the same input.

However, the approach also has some limitations. When solving the optimization problem, audit engagements can only be filled with the currently available workforce. For example, if the audit company accepts more audit engagements than it can handle with the current workforce, some audit tasks are left unassigned. Therefore, the lead auditor needs to check the final planning to make sure all necessary staff are available for the individual planned audits.

Another limitation could be the user-friendliness of the final output. According to the expert, it can be read by everyone and does not require familiarity with the machine learning mechanisms used. However, the depth and complexity of the results can be demanding, because the planning depicts a whole audit season with all audit engagements and all staff members.

Ongoing maintenance of the workforce planning tool is necessary as relevant regulations (e.g., work laws) change over time. Furthermore, once the audit season is finished, the audit staff plan is used to recalibrate the machine learning algorithms for the next round of staff planning. The whole process starts again every season, but the accumulation of more and better data and should yield continuous improvements over time.

Using a workforce planning tool in an audit company is outside the immediate scope of ISA. The individual audit engagement plan, which is required by ISA, is designed by the engagement leader and is also needed for the workforce planning tool. Using such a tool for reasons of convenience does not create any conflicts with the existing auditing standard, as long as it is ensured that no tasks for individual engagements are unstaffed when the audit is performed.

However, this use case is affected by ISQC 1, which regulates the quality control for audit firms. In particular, ISQC 1.31 requires that appropriate personnel be assigned to the audit engagements. Hence, the presented application relies on reliably maintained optimization criteria, especially regarding the entered level of experience. Therefore, keeping the input data current and accurate is critical.

Compared to other potential applications, this one is unique as it focuses on the needs of an entire audit company and not a single audit engagement. This generates some advantages, for example, concerning data quality, because audit companies use their own data, which is in their sphere of control. Moreover, focusing on the improvement of the processes of a whole audit company with data analytics is something the reviewed literature in chapter 4 has not yet picked up.

6.7 Summary

The future potential applications reviewed in this chapter are presented according to their areas of application in Figure 25 on the next page. The usual differentiation between conventional and advanced data analytics is not necessary as only advanced methods were showcased.

The presented applications show promise for the audit industry. In particular, the expert interviews suggest that machine learning methods can be applied within the audit process, even if this is not yet currently the case in the interviewed companies. However, the proposed applications first need to be tested in practice to confirm their usefulness for the auditor.

All presented applications should be regarded as only a supplement for the auditor, and in most cases, manually following-up on the results is required. By using very simple algorithms (clustering, association rules, and predictions) initial tests can be conducted at almost no additional cost (e.g., only requiring an Excel add-on) and with basic data analytics knowledge. However, for more complicated methods and models a data scientist is necessary.

Regardless of which data analytics tool the auditor uses, it is crucial, in order to correctly interpret the results, to have an idea of how the results of the data analytics were obtained and how it can be used for the audit. This is especially so for the first four described applications which directly provide input for the audit engagement itself. It is perhaps less critical in the final case of workforce planning.

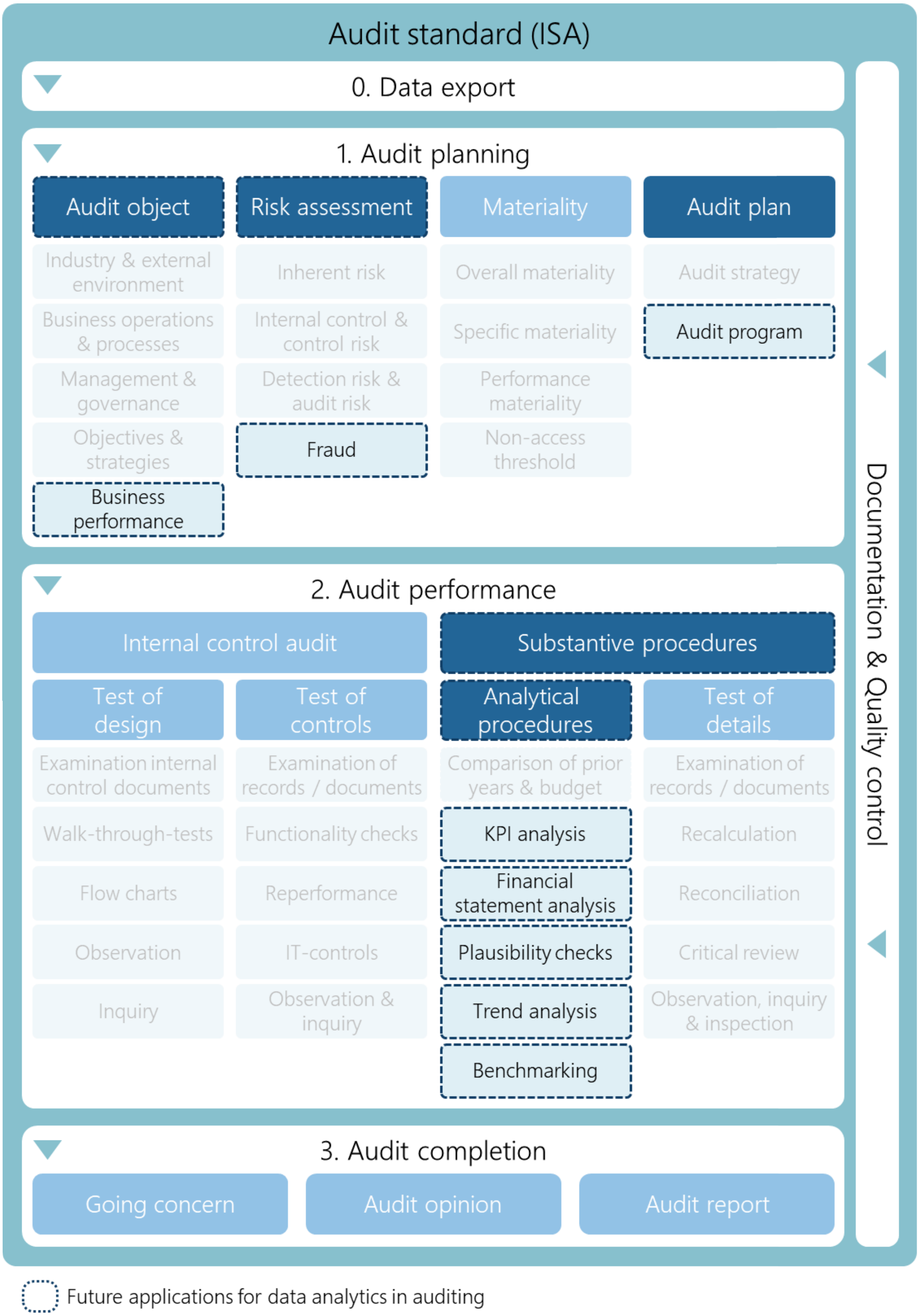

Figure 25 Illustration of future applications identified for data analytics in auditing

7 Implications for the audit market, audit profession, and regulator

7.1 Research methodology

The previous three chapters presented the various applications of data analytics in external auditing. This chapter focuses on the fourth research question:

What are the implications of using data analytics in external auditing for the audit market, audit profession, and regulator?

The implications drawn by the author are derived from data collected in previous chapters and validated with expert interviews.

7.1.1 Sample selection

A final round of focus interviews was executed to validate the implications. This could be described as an ex-ante evaluation (Sonnenberg & Vom Brocke, 2011), as the audit profession only recently started using more advanced data analytics tools, and the full implications are not yet realized.

For validation of the implications discerned, six interviews in total were conducted with:

- Swiss expert association for audit, tax, and fiduciary (EXPERTsuisse)
- Swiss expert association for fiduciary (TREUHAND|SUISSE)
- Swiss federal audit oversight authority (FAOA)
- Partner at a Big 4 audit company focused on data analytics
- Professor for auditing and internal control who is also a partner at a Big 4 company
- Owner of a small audit firm

These six experts were selected based on their publications about the future of auditing or their daily work in this area. The Swiss expert association for audit, tax, and fiduciary is the principal association for Swiss auditors. Smaller audit companies that offer audit and fiduciary services are also affiliated with the expert association for fiduciaries. To cover this segment of the audit market as well, an interview was conducted with this association and with a small audit company.

7.1.2 Validation procedure

For the validation procedure, the decision was made to carry out individual focus interviews to make sure each interviewee was able to reflect openly about the discussed topic and was not affected by the presence of others. Therefore, each interview was scheduled separately for one hour to foster an open discussion. All interviews except one were carried out in person. During the focus interviews, a bullet point list of the researcher's implications was presented and discussed with the participant. Depending on the interview partner, the order of the topics was changed to address the most relevant points for this expert first (e.g., the interview with FAOA started with the implications for the regulator).

A conscious decision was made between recording the interviews ensuring a complete documentation of the spoken content and having the opportunity to discuss the implications in an open discussion. The latter was given more importance and therefore, the interviews were not recorded[12] due to the confidential content and to ensure a barrier-free and honest discussion. Instead, notes were taken by hand during the interview.

[12] Not recording the interviews also eliminated the possibility for coding.

Right after each interview, both validation evidence for and counter-evidence against the initial implications identified by the author were captured in a template (Tremblay, Hevner, & Berndt, 2010), to mitigate the downside of not recording the interviews. Once all interviews were conducted, inputs from all interviews were reconciled and the conclusions were finalized.

The implications are presented separately for the audit market, audit profession and regulator.

7.2 Implications for the audit market

Based on the assumption that the implementation of data analytics will go beyond analytics for journal entry testing, future auditors will offer a broader service spectrum ranging from “traditional” to “data driven” audits. The decision power over which type of audit is chosen, as also over which auditor gets the assignment, lies in the hands of the audit clients. As auditors are expected to cater to the needs of each individual client, the future of the audit market depends heavily on the client decisions.

A recent study investigated the degree of digitization of Swiss businesses, which was found to be not particularly high at the moment (Klauser & Herzog, 2017). For a broad application of a data driven audit, the evolution on the client’s side is crucial in regards to digitalizing and standardizing their processes as much as possible. Otherwise, some tools will be limited in their usability. Therefore, an underlying hypothesis for the assumptions made here, is that more digital data will become available over time. In this journey, audit companies might be supported by the general market, as digitalization and standardization promise efficiency gains.

The digital availability of data is a prerequisite for any client requesting a data driven audit. Clients with ERP-systems are more likely to have the needed data, as more data is captured in such systems. One interviewed expert partly rejected this point, because the data could be combined from various systems (e.g., accounting system, customer relationship management) to provide a similar database as an ERP-system. However, as the auditor is under time pressure and combining the data from several source systems takes time and resources, it is unlikely that an auditor would undertake to build such a database himself unless the audit client is willing to pay for the effort.

Nevertheless, the auditor will already have access to all journal entries of a given year, which is the complete backbone of the financial statement, as testing these journal entries is obligatory by the current ISA standard. Therefore, at least this dataset will always be available, no matter what kind of audit is chosen, and the auditor will be able to apply whatever data analytic tools he chooses for journal entry testing.

7.2.1 Competition

Data analytics is currently primarily the domain of larger audit companies. This is supported by the following indicators: Larger audit companies are already increasingly asked to include data analytics in their request for proposals (Schneider & Bradtke, 2017). They are also investing heavily in this domain (e.g., Deloitte, 2016; Ernst & Young, 2018; KPMG, 2017; PwC, 2019). By contrast, a non-Big 4 audit company declined an interview due to the limited relevance of the topic for them. The interviewed expert from a smaller audit company for this chapter confirmed this as well. A quote from the novelist William Gibson (1999) neatly captures the situation: "The future is already here — it is just not very evenly distributed".

As the Big 4 companies strive for growth and have more tools currently available for performing data analytics, it is assumed that these companies will use these tools to try to get ahead in their fierce market competition with each other. However, if there is a technological superiority of one audit firm in the market, it is presumed to last only for a short time. Once awareness spreads that one player has found and implemented a successful technology, competitors will copy their approach as quickly as possible in order to catch up (Harvard Business Review, 2012). Furthermore, it is common for audit professionals to switch between audit firms, further promoting the spread of new technologies. In the case that a useful technology is developed outside of an audit company (e.g., by a startup), it will be even more accessible for other players since the startup will attempt to sell it to as many clients as possible.

Another aspect of the competition scenario is potential "lock-in" effects as some newly developed audit tools entail set-up costs to calibrate them to the individual audit client. Once these costs are amortized, the current audit company can offer an audit at lower fees than competitors who still have to price the initial set-up costs into their offers. The only way for competitors to break this "lock-in" effect is to sacrifice profits in exchange for a better chance of winning the mandate. Such strategies have the disadvantage of prolonging the amortization phase of the set-up costs, which also has an effect on the amortization of the developed tools.

Three out of six experts said that market competition for large audit engagements is so fierce, that audit companies willingly sacrifice profits in the first year to gain a prestigious audit client, while amortizations over time are already expected. If this behavior becomes more pronounced due to the implementation of data analytics, then it will have a market-wide impact of compressing margins.

Patterns of disruptive innovation observed in other industries challenge assumptions that innovation will come from the Big 4; there may well be opportunities for new technology-driven audit startups. In the medium-term, however, the appearance of a true "disruptor" for the audit market seems unlikely on a global level, as huge entry barriers exist. The history of the Big 4 companies shows that they have emerged from various mergers and acquisitions. Once this purchase phase was completed, they enforced a global standard ensuring that audits conducted within their firm network yield a certain standard worldwide.

Interestingly, two experts saw banks as potential local disruptors, because some now offer the service of creating a client's financial statements if they only hold accounts at the one bank. The next step could be that such banks expand their offer to auditing or reviewing the statements (limited statutory examination), a move which would obviously impact the audit market.

7.2.2 Adoption lifecycle

The question is when the roll out of advanced data analytics will be applied in audit engagements, leading to the broad adoption of data driven audits. To analyze this question further, the rough segmentation of audit clients developed in chapter 5, into "insight clients" and "price-sensitive clients", is extended. The technology adoption lifecycle from Moore and McKenna (1999) (Figure 26) is used to present the scenarios.

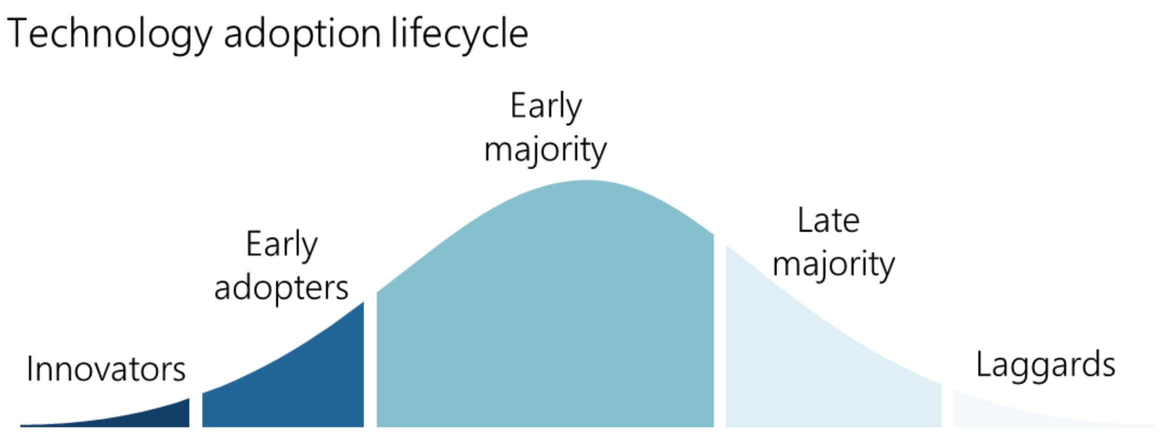

Figure 26 Technology adoption lifecycle based on Moore and McKenna (1999)

“Insight clients” want more than just the audit opinion from the auditor; they also want knowledge about their company, which is provided by using data analytics during the audit. Therefore, they are "early adopters" (Moore & McKenna, 1999) of data driven audits. Thus, this client segment encourages and supports investments in new tools from an early stage. This is the client segment most willing to support these innovations financially through higher audit fees. These clients create the incentives and nudge auditors to provide data driven audits.

“Insight clients” already receive more knowledge through the use of process mining in audits today. According to two experts, especially larger companies with audit committees and the respective legal duties are more interested in options that provide them with more insights. In general, it is presumed that competition within the audit market to gain “insight clients” will be intense.

“Price-sensitive clients” will probably at first avoid data driven audits and instead prefer a traditional audit process, as the former will be initially more expensive for them. They will stand on the sidelines while audit firms, together with the “early adopters”, tackle the various hurdles that need to be overcome to make data driven audits more routine.

According to one expert, the current simplified cost calculation formula for audit fees is “hourly fee x human hours required for the audit”. For data driven audits, a new component will be introduced for “technology costs”. For a data driven audit to directly compete with a traditional audit in the world of “price-sensitive clients”, the “human” hours required for an audit must be reduced correspondingly to offset these additional technology costs. It is assumed that only once data driven audits are more routine for the audit companies, the required efficiency gains necessary for this can be realized. Only when the audit fee of a data driven audit is lower than the audit fee of a comparable traditional audit will “price-sensitive clients“ have an incentive to switch to a data driven audit. This scenario holds under the assumption that the price reduction is not at the expense of more work on the client's side. Thus, "price-sensitive customers" will most probably follow a typical "late majority" behavior pattern in the technology adoption lifecycle (Moore & McKenna, 1999).

A scenario raised by one expert was that the auditor could also act as an enabler, by showing clients what insights can be gained with advanced data analytics. Thus, some "price-sensitive" clients might convert to become "insight clients" and be willing to pay for the use of advanced data analytics during an audit. These clients would become part of the "early majority" phase of the technology adoption lifecycle.

It cannot be completely ruled out that a need for traditional audits will always exist. They will, at least, probably continue to exist for a long time, as predicted for the "laggards" group in the technology adoption lifecycle (Moore & McKenna, 1999).

7.2.3 Small audit companies

In Switzerland, the special case of a limited statutory examination exists. Such audits are generally outside of the scope of this dissertation. However, small audit companies are an important part of the audit market, and this subchapter is focused on them.

Based on one expert interview, small audit companies are already performing conventional data analytics methods (e.g., rule-based queries). However, when it comes to advanced data analytics, it is questionable if the potential efficiency and quality gains from advanced data analytics outweigh the set-up costs. This is primarily because limited statutory examinations have a smaller scope, and therefore also a lower audit fee. Based on the expert's opinion, it will be a trade-off between how much work is "taken" from the auditor's plate through process improvements and how much is "added" with newly required steps. Even in a digital society, small companies will produce only a small amount of data as they do not perform high volumes of transactions. It is likely that the traditional audit approach will survive the longest in this market segment.

The Swiss CPA education now includes some data analytics courses. Therefore, eventually even smaller audit companies that would like to implement more data analytics in their audits will have an opportunity to do so through knowledge brought in by new hires. However, one expert feared that smaller audit companies will have limited access to data analytics-savvy auditors in the foreseeable future as they will be in high demand.

Smaller companies with a naturally smaller audit scope will at least initially not generate enough income from audit fees to carry the added costs of using advanced data analytics. Furthermore, they could lack the data volumes required for the use of advanced data analytics methods.

Two experts expected that small, technologically innovative audit companies will arise. It seems likely that smaller audit companies will cater to the need for data driven audits of smaller, but digitalized, companies, or will develop the capabilities to perform “innovative” forms of traditional audits. To make this possible, providers offering tools for smaller audit companies will be necessary. Such providers already exist on the market today and are expected to evolve. They may start by giving smaller audit companies access to innovations which they cannot develop in-house. Being able to buy software licenses could become part of the entry barrier to the market in a data driven audit world.

7.3 Implications for the audit profession

Under the assumption that the audit market in the future will demand data driven audits, the following implications apply for the audit profession.

7.3.1 Extension of skillset

To make sure that auditors can provide data driven audits, changes in education and training need to take place. The Swiss expert association for audit, tax, and fiduciary has already reacted to this anticipated demand and just recently added a data analytics course to the Swiss CPA education. Data analytics courses are also offered to already licensed auditors. Larger audit companies offer in-house training courses to make sure their employees are trained adequately.

Given the wide array of challenges that currently exist and have been discussed in subchapter 5.2.5 (e.g., challenges at the client and within the audit company itself), learning on the job will be necessary to strengthen and increase the application of the developed data analytics tools during the audit engagement. In addition, the fast-developing nature of advanced data analytics technologies, will require ongoing training. Over time, this will raise trust within the audit community towards such applications, and confidence in using them.

Universities play a key role in terms of what kind of education is offered during undergraduate and graduate studies for future auditors. One expert supported data analytics becoming a compulsory subject, as many professions beyond auditing benefit if employees have such knowledge. Weissenberger, Förster, Bravidor, and Wesser (2019) conclude that the auditor also needs more IT-know-how, to fully understand the processes of a company and how they are reflected in the system.

Thus, data analytics is becoming a necessary extension of the auditor’s skillset. It will be required to ensure the auditor is capable of applying professional judgment and skepticism, of questioning and critically evaluating the results of data analytics.

Being familiar with data analytics also helps the auditor to be aware of the potential limits of data analytics methods. First, the auditor needs to focus on the data itself (quality, attributes, creation), which means having IT-reliance for the individual client will play a crucial role. Second, in a "digital world," incentives are provided to commit fraud "offline" or more effectively with the help of machine learning (e.g., Schreyer et al., 2019a). As is the case when fraud is concerned today, the auditor needs to be aware of what changes and challenges can occur.

Introducing data analytics in auditing also influences the timely aspect of an audit. As discussed in chapter 5, a new stage "data export" is added at the beginning of the audit process. This could also lead to starting an audit earlier, as the new stage requires some time to conduct. In a data driven audit more weight will be given to the audit planning stage as the data will already be examined in this stage. Therefore, more audit work is pushed into the audit planning phase. Moreover, it can also lead to taking away some work in the audit performance stage as the knowledge accumulated in the audit planning stage can be used for the former.

7.3.2 Dealing with increased transparency

Especially with the more advanced methods, increased transparency is achieved when analyzing the audit object, which can lead to a better audit quality (IAASB, 2016). This can transform the auditor's experience. Where previously he saw small bits and pieces due to the sampling method, now he is confronted with a completely transparent picture produced by analyzing all available data.

In some cases, this means opening a pandora's box and revealing unpleasant facts previously unseen. When such revelations occur, the auditor needs to address them adequately and might be forced to do more follow-ups than with the traditional sampling method. However, this could trigger negative consequences in the absence of appropriate standard guidance. Situations could be created where auditors, in their own and their client's best interests, avoid asking specific questions of the data to circumvent the need to deal with the answers.

7.3.3 Industrialization of auditing

Currently, and in the foreseeable future, data analytics is only a supporting tool for the auditor. To change this, new or existing technologies still need to evolve. Three experts felt that even the advanced data analytics technologies that currently exist are not yet ready for practical adoption and still need to evolve further. However, one expert objected to this perspective, saying auditors should take advantage of all opportunities provided by technologies today.

With technological progress and the gradual change to data driven audits, it is assumed there will be an increase in the number of data analytics tools in the auditor's toolbox. Tools that begin by supporting auditors in performing parts of individual audit steps may in time be capable of delivering full audit steps (automatization). This opens a path for a step by step evolution from today's manual approach to auditing to an industrialization of the process in the longer term.

As new technology continues to develop, new revenue opportunities arise, according to the input of two experts. One possibility is for Big 4 companies to offer the developed tools on the market to smaller audit companies. This would make sure the amortization of their investments is reached faster and new revenue streams are created. However, it would also mean directly supporting the competition, which could be problematic.

Any fear auditors may have of being replaced by artificial intelligence seems unreasonable, at least in the medium-term and because they are protected by the audit standard. However, by consequence of industrialization, it could lead to further consolidation of audit firms in the market, as has happened in other industries. One expert added that already today a general trend towards consolidation in the sector of small and medium audit companies can be observed.

The auditing industry has been using conventional data analytics methods (e.g., rule-based queries, descriptive statistics) for some time. However, when it comes to implementing advanced data analytic tools, audit companies are just getting started. This has been confirmed by the recent study of EXPERTsuisse, which revealed that only 22% of Swiss auditors are using data analytics software (Klauser, 2019). This is surprisingly low, given the current hype. While new technologies such as machine learning have been enjoying early successes in other areas, external auditing has been quite resistant to change.

An interesting input from one expert concerned how long the implementations of newly developed data analytics tools take in a global audit company. Based on his experience, a worldwide roll out of a new tool can take up to eight years. First tests are piloted in one local office, followed by a few local offices and with a selected number of audit clients. If this is successful, the tools are then slowly rolled out to all local network firms, each of which then goes through its own sequence of adoption phases.

This suggests that audit companies of global networks collectively may be destined to be late adopters of new technologies, compared to companies from other industries, as it clearly takes a long time for the developed tools to be fully deployed. However, as larger audit companies are more occupied with the matter of data analytics today, they might be the first among what is generally an industry of late adopters. In a rapidly developing world, eight years can be a long time. Even though the slow deployment processes of large companies could give an advantage to smaller audit companies, the latter probably do not have the resources to afford investments in technological developments.

Overall data analytics gives the potential to industrialize external auditing. However, as shown this will probably be a slow process with small increments at a time.

7.3.4 Transformation of the pyramid system

As more data analytics tools arrive in the auditor's toolbox, the question is: who will apply these tools? As shown, performing advanced data analytics methods is a demanding task. Even though the user-friendliness of the tools is improving, the technical expertise of a data scientist is usually still necessary for performing advanced data analytics methods (Earley, 2015). This is especially important to avoid the erroneous use of such methods (Kelleher & Tierney, 2018).

Therefore, for a data driven audit, and especially when using advanced data analytics, it should be expected that the circle of experts on which the auditor relies on will grow to include data scientists. This could mean that auditors in the future will have a data scientist in addition to IT-auditors (Küng et al., 2017), which affects using the work of an "auditor's expert" (ISA 620).

The majority of experts generally agreed with this assessment. However, one expert objected because in his opinion, such tools need to be built in a way that the auditor can use them himself, and hence, the usability of these tools must improve. Otherwise, it will become harder for the auditor to perform judgment.

Additionally, it could be challenging in an industry under price-pressure to add such a high-priced labor component (data scientist) to the service. A skilled workforce comes with high salaries, which are contradicted by the contemporary downward pressure on audit fees. One solution mentioned by Salijeni et al. (2019) is to outsource such work to countries with high IT-expertise and low labor costs.

The process of auditing includes many mundane tasks (Issa et al., 2016). Standardized, repetitive activities, currently carried out by assistants or in shared service centers, could be simplified or carried out by machines (e.g., natural language processing). How will the assistants' work and shared service centers develop on a long-term basis? Several scenarios are possible and two are addressed here briefly:

Assistant's work becomes more demanding; need for shared service centers is reduced

Based on the assumption that assistants will still be required to process the information gained from the tools and to prepare the audit documentation for the review of senior staff, a complete abolition of this group of employees is unlikely. While the work of assistants may be subject to partial automation, the work which is left, such as following up with clients to discuss identified deviations, is more demanding. Thus, it is expected that fewer but more qualified assistants will be required to conduct an audit.

Meanwhile, shared service centers will likely be less needed, as the work they perform will be increasingly automated.

Insourcing of shared service centers and changed career start of assistants

Another possibility raised by an expert was the insourcing of shared service centers, meaning that previously outsourced tasks are brought back from foreign countries to the country of the audit company, where the assistants' work will be bundled in such service centers. As a consequence, this could mean that every new assistant starts in a service center, and service centers become the nurseries where auditors start their careers.

If many routine tasks are automated or at least supported with data analytics applications, while at the same time the requirements for expert judgment increase, the question arises where the next generation of senior auditors will come from. All of these factors – new auditor's experts, more skilled workforce required, change of assistant's work – may be incompatible with the current hierarchical pyramid structure of audit companies. One tempting response would be to cut costs by narrowing the base of the pyramid. At an industry level, however, this has the potential to create a severe pipeline problem over time, as fewer assistants will result in a smaller pool of candidates for the next generation of senior auditors. These issues could be exacerbated by different job expectations that millennials have (Bannon, Ford, & Meltzer, 2011).

Overall, advanced data analytics technologies will perhaps have a much broader organizational impact on audit companies than is immediately apparent.

7.4 Implications for the regulator

Both developments in the audit market and in the audit profession have an impact on the regulator, which is the oversight authority.

As first noted in subchapter 5.2.6, the application of data analytics in external auditing is not governed explicitly by the standard setter at the moment. As a consequence, and until this changes, the evaluation of the correct application of data analytics in auditing is left to the regulator. The Swiss regulator, for now, is taking an observatory stance, going by the materials on data analytics published in the annual reports (FAOA, 2017, 2018, 2019a).

The question arises whether the work of the regulator will change if more advanced data analytics methods are applied in audits in the future. The current ISA standard requires that audit documentation be written in such a way that an independent and experienced auditor can follow the decision-making process (ISA 230). This places responsibility on the auditor in charge to make sure that the documentation is understandable for auditors, which puts the regulator in a comfortable position to evaluate the application of data analytics. Future developments will show if a data scientist will be necessary also on the regulator's side, to review the evidence gathered by data scientists on the audit company's side.

7.4.1 Predicament for the regulator

In a regulated world, such as the one the auditor operates in, an unregulated state is unusual and can cut both ways. Actively applying data analytics requires the auditor to close potential gaps in the audit standard himself, which requires extra resources. However, not applying data analytics and just auditing the required sample can protect the auditor from doing more work or cause him to refrain entirely from performing data analytics. Therefore, the regulator could potentially play a vital role in fostering the use of data analytics methods within the industry. First, by providing the necessary openness to allow such methods to be applied. Second, by "encouraging" audit companies to use such tools under the assumption that a higher audit quality is provided (IAASB, 2016) as various stakeholders can benefit from a higher audit quality, including the regulator.

However, "encouraging" the auditor is not a trivial endeavor, especially given the limited room for trial and error that exists in the heavily regulated and litigious environment that he operates in (Alles & Gray, 2016). For example, when a salesperson implements neural networks to predict if costumers will churn, the downside is limited, while the upside is potentially considerable since valuable knowledge on avoiding churn can be gained. The situation looks entirely different for the auditor: the upside is limited, and the downside is potentially massive. As the auditor is responsible for the audit opinion – with his license and career on the line – he will think twice before trusting and implementing new technologies for which no guidance is yet provided. Given this incentive structure, the cost-benefit analysis for implementing advanced data analytics applications, is different for audit companies when compared to other corporations (Alles & Gray, 2016).

On the one hand, the regulator could encourage the application of data analytics in external auditing. On the other hand, the regulator is responsible to oversee the adequate use of data analytics, which creates a predicament for the regulator.

7.4.2 Ensuring responsible use of data analytics

Even when auditors do shoulder the burden of self-regulation where the standards are silent, the regulator still needs to be the critical judge and make sure that the auditors apply the tools in line with the current auditing standard. As each audit company develops its internal guidance for applying data analytics, the regulator needs to check the internal guidance. In the case of international audit companies, which follow the same ISA guidance across several countries, local regulators need to work together to agree if the internal guidance is appropriate, according to one expert.

As the tools become better and generally more accurate, a new risk will be overreliance on these tools. In general, it is assumed that failure will over time becomes less frequent, but inevitably, it will happen.

Hail, Tahoun, and Wang (2018) find that governing standards are launched after the first major scandal occurs. Based on this finding, it depends on what kind of scandal comes "first", regulations will then be created for it, and will influence the future of data analytics in auditing.

The following two scenarios may be considered:

Scenario 1: Data analytics are not applied, and material misstatements are not identified. If data analytics had been applied, the misstatements could have been detected with ease. This could create a call for the use of data analytics to become mandatory. To enforce this, the standard setter would need to explicitly specify which methods must be used.

Scenario 2: Data analytics are applied, and material misstatements are not identified. As a result, strict regulations that make the use of data analytics more difficult may be put in place.

In either scenario, specific regulation for using data analytics in auditing is the result. This has the benefit of removing uncertainty for those audit companies which apply data analytics. However, in the first scenario, it may create challenges for audit companies not already applying data analytics that are then forced to get familiar with the topic quickly. In the second scenario, there is a significant risk of overregulation for data analytics with a negative effect on innovations in auditing.

8 Conclusion

In the following subchapters, the findings of the research questions will be summarized, followed by limitations and an outlook for future research.

8.1 Summary of findings

This dissertation opened with the disruptive potential of data analytics covering technologies such as artificial intelligence in the audit profession. Findings suggest that the impact will be great, but that the artificial intelligence auditor of the future is still beyond the visible horizon. Therefore, the anticipated disruption is not yet discernable or at least still lies in the more distant future. This is primarily due to the following factors. First, the effective application of advanced data analytics technologies (e.g., machine learning) in external audit practice, is still in its early days. Second, the current applications are just supporting tools for the auditor. Third, a variety of hurdles have to be overcome first (e.g., client concerns, hurdles within the audit companies, data extraction, data security, data storage). Fourth, the auditor is protected by the current audit standard as data analytics alone does not provide adequate audit evidence to obtain an audit opinion. However, data analytics do have a wide range of applications in the external auditing process.

The first research question focused on identifying use cases in research, and delivers initial evidence of this: 42 use cases in total were identified, split among the fields of fraud (23), going concern (8), analytical procedures (3), audit opinion (3), risk assessment (3), and internal controls (2). The majority of these papers applied machine learning methods (33), while a smaller number used process mining (4), natural language processing (3), and rule-based analysis (2). The machine learning use cases depicted in the literature still need to evolve before becoming practical for auditors. However, process mining, which emerged first in research, is already known among practitioners.

This leads to the second research question, which focused on the current application of data analytics in practice. Unsurprisingly, conventional data analytics methods (rule-based methods, visualization, and descriptive statistics) are widely used and have been used for a while in the audit profession. Few companies apply advanced data analytic methods (process mining, machine learning, and natural language processing) as of today. Especially the interviews with audit companies revealed a wide range of current use cases for data analytics in auditing. While the majority of identified applications today are conventional methods, they prepare the ground for future applications of advanced data analytics methods.

This was the focus of the third research question, which investigated potential future applications. Among advanced data analytics methods, machine learning methods are particularly rarely implemented in the audit process today. Therefore, the primary focus was to establish realistic future potential applications for these methods. Five future potential uses cases were uncovered. Three of them focus on journal entry testing, including the possible detection of fraud and identifying material misstatements. One looks into forming predictions for analytical procedures and another one into workforce planning for a whole audit company.

When all changes due to the use of data analytics and all identified use cases are mapped to the audit process, the following data driven audit process emerges (Figure 27).

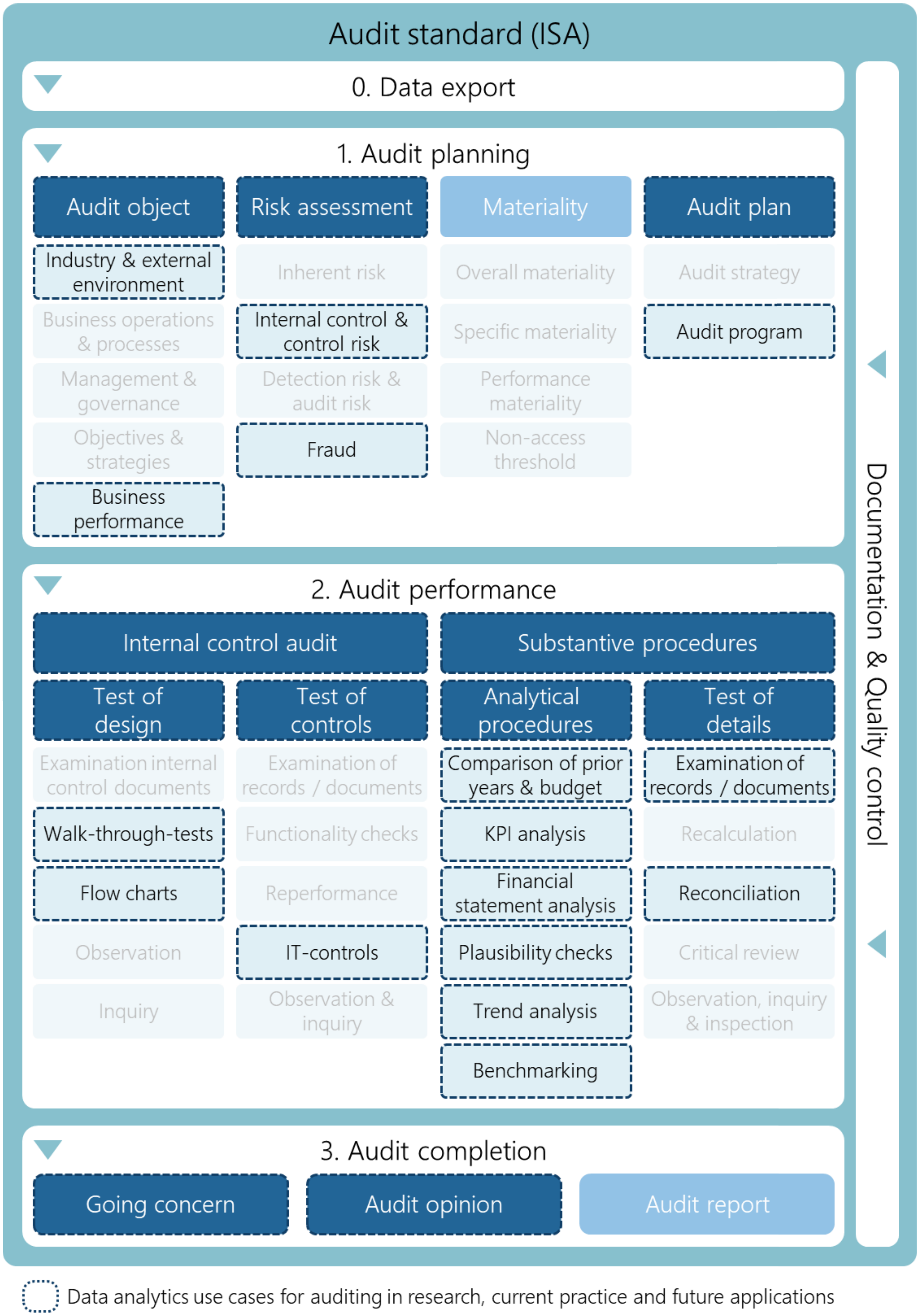

Figure 27 Illustration identified use cases for external auditing

It is evident that data analytics can be applied to many parts of the auditing process. However, the anticipated increase in the use of data analytics in the audit profession comes with consequences. This was the focus of the last research question, which investigated the implications for the audit market, audit profession, and regulator.

Regarding the implications for the audit market, it is assumed that the auditor depends on the desire of the audit client for a “data driven audit”. Auditing itself is a well-defined service offered by many audit firms in a competitive market. The end product of this service, the audit report, is highly standardized. What is offered depends on what the client desires. Therefore, the evolution of the audit market depends heavily on the needs of the clients. However, the broad adoption in the audit industry of data driven audits will most likely take place in small increments, as clients see its benefits.

Ultimately, all professions evolve over time – the auditor of the past is different from the auditor of today, who will, in turn, be different from the auditor of the future. It is fair to conclude that data analytics will play an increasing role in this profession, further expanding the required skillset of an auditor. Transformation and change will take place and will likely also influence the organizational set-up of audit firms. As more data analytics techniques are implemented in auditing, this could lead to an industrialization of the audit profession.

Since the current ISA standard does not address data analytics, both auditors and regulators are without practical guidance. Currently, the regulator acts only as an observer in regards to the expanding application of data analytics in auditing (FAOA, 2019a). In the future the regulator will have to find a balanced approach, on one side encouraging the application of data analytics as it promises to increase audit quality (IAASB, 2016), while on the other side still fulfilling its duty to ensure the adequate application of the audit standard.

In general, all findings of the research questions contribute to the ongoing discussion of this topic and provide new insights. With the four research questions addressed, a holistic overview of data analytics usage in external auditing is provided.

To sum up, this dissertation sheds light on the applicability of data analytics in external auditing, especially for practitioners, and its implications on a broader scale. In addition, this work actively contributes to the academic research by answering calls for research (Appelbaum et al., 2018; Earley, 2015; Issa et al., 2016; Wang & Cuthbertson, 2014).

8.2 Limitations

This dissertation and its results are subject to limitations. As for any academic work, a scope had to be defined. A particular challenge is the blurry and broad definition of data analytics and the multitude of methods it encompasses. It would be practically impossible to cover the full spectrum, and the author had to restrict herself to a carefully curated selection.

In addition, it was necessary for this dissertation to present auditing as a single standardized process. To do so, it was necessary to abstract the whole process and "simplify" it to some degree. Even if this simplification is a correct representation of a regular external audit, it does not adequately reflect the whole complexity of an audit in the real world. It also neglects special audit cases (e.g., first-time audits, group audits), industry-specific variations (e.g., banks or insurance companies), and the additional requirements that apply to publicly listed companies.

It was also necessary to constrain the scope of the literature review. It would have been interesting to also include the application of data analytics in other industries.

Due to the exploratory nature of the research design, the results cannot claim comprehensiveness. Although appropriate due diligence was exercised and the use cases were examined from three different angles (research, audit practitioners and data analytics experts), the discussed use cases should not be interpreted as a complete representation.

The usual caveats around the subjectivity of interviewees and local restrictions apply. In addition, the future potential applications discussed in chapter 6 still have to be tested in a real-world audit setting. As more experience is gained, practice will show some of them to work, while others will not prove to be as valuable as anticipated and will subsequently be discarded.

Finally, for such a popular topic at the interface between the present and future, there are, by nature, uncertainties. The implications discussed in Chapter 7 constitute statements about the future, which are challenging to get right and by definition, never complete.

Many of the presented limitations of the dissertation provide opportunities for future research.

8.3 Outlook for future research

Since this dissertation covers a broad field, plenty of further research directions are open. The following overview focuses on possible extensions of this dissertation.

Even though the IAASB (2016) mentions that data analytics could enhance audit quality, no actual proof exists of such an improvement, to the researcher's knowledge. Therefore, future research could investigate if this can be shown to be the case.

Successful use cases for data analytics in related industries (e.g., internal audit, management accounting) or other industries (e.g., information management, business innovation) could be examined for their potential transposition to the audit process.

As the selection of audit companies focused on medium-sized and larger audit companies, a similar research design could be applied to smaller audit companies in Switzerland. Another opportunity would be to apply the same research design to audit companies in other countries and compare the results.

This work focused exclusively on data analytics. Other emerging new technologies – such as blockchain and robotic process automation – would also be exciting fields for research.

Since some future potential applications presented in chapter 6 have not been tested with real-world data in a practical context, future research could investigate the realization of these applications.

The main objective of this dissertation was to establish the applicability of data analytics, and especially a subset of selected methods. Further research is needed to go into more detail on specific applications and methods and to evaluate them in a practical use context. More research is also needed on specific parts of the data analytics lifecycle (e.g., preprocessing, deployment). Research into the process of successfully deploying new tools in practice and ensuring their correct use by auditors would be both a valuable and practical contribution.

Under the hypothesis that future audit clients will be able to choose between traditional or data driven audits, an interesting research field would be whether tech-savvy audit clients or audit committees are more likely to request data driven audits earlier than other companies.

Finally, data analytics in the audit process has the potential to enable a continuous auditing process. This could be a highly transformative change that could also be addressed by future research.

At any rate, as the author of this dissertation has found on her journey, the future of data analytics in auditing is an exciting and promising field for any researcher. It is one that will yield much fruit over the coming decade as an established and respected profession reinvents itself for the digital age.

References

Abbasi, A., Albrecht, C., Vance, A., & Hansen, J. (2012). Metafraud: a meta-learning framework for detecting financial fraud. *MIS Quarterly, 36*(4), 1293-1327.

Abbott, D. (2014). *Applied predictive analytics: Principles and techniques for the professional data analyst*. New York, NY: John Wiley & Sons, Incorporated.

Abts, D., & Mülder, W. (2017). *Grundkurs Wirtschaftsinformatik. Eine kompakte und praxisorientierte Einführung* (9th ed.). Wiesbaden, Germany: Springer Fachmedien

ACFE. (2019). Report to the nations 2018: Global study on occupational fraud and abuse. Retrieved July 29, 2019, from https://s3-us-west-2.amazonaws.com/acfepublic/2018-report-to-the-nations.pdf

Aggarwal, C. C. (2015). *Data mining: The textbook*. Berlin, Germany: Springer International Publishing.

Aggarwal, C. C. (2017). *Outlier analysis* (2nd ed.). Berlin, Germany: Springer International Publishing.

Aggarwal, C. C., & Han, J. (2014). *Frequent pattern mining*. Berlin, Germany: Springer International Publishing.

Albrecht, W. S., Albrecht, C. C., & Albrecht, C. O. (2004). Fraud and corporate executives: Agency, stewardship, and broken trust. *Journal of Forensic Accounting, 5*, 109-130.

Alles, M., & Gray, G. L. (2016). Incorporating big data in audits: Identifying inhibitors and a research agenda to address those inhibitors. *International Journal of Accounting Information Systems, 22*, 44-59.

Alles, M., Kogan, A., Vasarhelyi, M. A., & Brennan, G. (2006). Continuous monitoring of business process controls: A pilot implementation of a continuous auditing system at Siemens. *International Journal of Accounting Information Systems, 7*(2), 137-161.

Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of big data by the audit profession. *Accounting Horizons, 29*, 439-449.

Alvesson, M. (2011). *Interpreting interviews*. London, United Kingdom: Sage.

Anandarajan, M., Picheng, L., & Anandarajan, A. (2001). Bankruptcy prediction of financially stressed firms: An examination of the predictive accuracy of artificial neural networks. *International Journal of Intelligent Systems in Accounting Finance & Management, 10*(2), 69-81.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory, 36*(4), 1-27.

Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2018). Analytical procedures in external auditing: A comprehensive literature survey and framework for external audit analytics. *Journal of Accounting Literature, 40*, 83-101.

Arens, A. A., Elder, R. J., Beasley, M. S., & Hogan, C. E. (2017). *Auditing and assurance services: An integrated approach.* (16th global ed.). Boston, MA: Pearson.

Arnold, V., Collier, P. A., Leech, S. A., & Sutton, S. G. (2004). Impact of intelligent decision aids on expert and novice decision-makers' judgments. *Accounting & Finance, 44*(1), 1-26.

Babenko, B. (2008). Multiple instance learning: Algorithms and applications. Retrieved September 30, 2019, from http://vision.ucsd.edu/~bbabenko/data/bbabenko_re.pdf

Bannon, S., Ford, K., & Meltzer, L. (2011). Understanding millennials in the workplace. *The CPA Journal, 81*(11), 61.

Bartmann, C., Hufgard, A., Streller, S., & Weltner, V. (2018). Digitaler Aubrauch in der Wirtschaftsprüfung und Beratung. In Deggendorfer Forum zur digitalen Datenanalyse e.V. (Ed.), *Digitalisierung der Prüfung - Datenanalyse im Aufbrauch* (pp. 25-40). Berlin, Germany: Erich Schmidt Verlag.

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society, 78*(4), 551-572.

Blumberg, B., Cooper, D. R., & Schindler, P. S. (2008). *Business research methods.* (2nd European ed.). New York, NY: McGraw-Hill Higher Education.

Bogner, A., & Menz, W. (2009). The theory-generating expert Interview: Epistemological interest, forms of knowledge, interaction. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing Experts* (pp. 43-80). Basingstoke, United Kingdom: Palgrave Macmillan.

Bologna, G. J., & Lindquist, R. (1995). *Fraud auditing and forensic accounting* (2nd ed.). Hoboken, NJ: John Wiley and Sons.

Bönner, A., Riedl, M., & Wenig, S. (2011). *Digitale SAP®-Massendatenanalyse: Risiken erkennen-Prozesse optimieren*. Berlin, Germany: Erich Schmidt Verlag.

Brazdil, P., Giraud-Carrier, C., Soares, C., & Vilalta, R. (2009). *Metalearning: Applications to data mining*. Berlin, Germany: Springer.

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). *LOF: Identifying density-based local outliers*. Paper presented at the ACM sigmod record.

Brewerton, P., & Millward, L. (2001). *Organizational research methods : A guide for students and researchers*. London, United Kingdom: SAGE.

Burrus, D. (2018). A.I., machine learning, cognitive computing as profession's top technology hard trends for 2018: Part 1. *CPA Practive Advisor.* Retrieved September 12, 2019, from https://www.cpapracticeadvisor.com/accounting-audit/article/12410132/ai-machine-learning-cognitive-computing-as-professions-top-technology-hard-trends-for-2018-part-1

CaseWare. (2019). Yearly financial ratios of Microsoft. Retrieved October 18, 2019, from https://ai.caseware.com/company/0000789019

CaseWare. (n.d.). FAQ. Retrieved October 18, 2019, from https://www.caseware.com/artificial-intelligence-business-insights#faq

Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science, 56*(7), 1146-1160.

Cerny, P. A., & Proximity, M. A. (2001). Data mining and neural networks from a commercial perspective. Retrieved September 27, 2019, from http://www.orsnz.org.nz/conf36/papers/Cerny.pdf

Chamoni, P., & Gluchowski, P. (2017). Business analytics: State of the art. *Controlling & Management Review, 61(4)*, 8-17.

Chapman, P., Clinton, J., Kerber, R., Reinartz, T., Shearer, C., & Wirth, R. (1999). *Cross industry standard process for data mining (crisp-dm): Step by step data mining guide*.

Chen, H., Huang, S., & Kuo, C. (2009). Using the artificial neural network to predict fraud litigation: Some empirical evidence from emerging markets. *Expert Systems with Applications, 36*(2), 1478-1484.

Cleary, R., & Thibodeau, J. C. (2005). Applying digital analysis using Benford's law to detect fraud: The dangers of type I errors. *Auditing: A Journal of Practice & Theory, 24*(1), 77-81.

Committee of Sponsoring Organizations of the Treadway Commission. (2013). *Internal control – Integrated framework: Framework and appendices*. Jersey City, NJ: COSO.

Conway, D., & White, J. (2012). *Machine learning for hackers*. Sebastopol, CA: O'Reilly Media, Inc.

Davenport, T. H. (2014). *Big Data @ Work. Chancen erkennen, Risiken verstehen*. München, Germany: Vahlen.

Dbouk, B., & Zaarour, I. (2017). Towards a machine learning approach for earnings manipulation detection. *Asian Journal of Business & Accounting, 10*(2), 215-251.

Dean, J., & Whyte, W. (1979). How do you know if the informant is telling the truth? In J. Bynner & K. Stribley (Eds.), *Social Research: Principals and Procedures* (pp. 179-188). New York, NY: Longman.

Debreceny, R. S., & Gray, G. L. (2010). Data mining journal entries for fraud detection: An exploratory study. *International Journal of Accounting Information Systems, 11*(3), 157-181.

Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research, 28*(1), 17-82.

Deloitte. (2016). The power of advanced audit analytics everywhere analytics. Retrieved October 26, 2019, from https://www2.deloitte.com/content/dam/Deloitte/us/Documents/deloitte-analytics/us-da-advanced-audit-analytics.pdf

Dickey, G., Blanke, S., & Seaton, L. (2019). Machine learning in auditing: Current and future applications. Retrieved September 17, 2019, from https://www.cpajournal.com/2019/06/19/machine-learning-in-auditing/

Dikmen, B., & Küçükkocaoğlu, G. (2010). The detection of earnings manipulation: the three-phase cutting plane algorithm using mathematical programming. *Journal of Forecasting, 29*(5), 442-466.

Donnelly, L. (2018). Robots are better than doctors at diagnosing some cancers, major study finds. *The Telegraph.* Retrieved October 25 2019, from https://www.telegraph.co.uk/news/2018/05/28/robots-better-doctors-diagnosing-cancers-major-study-finds/

Doumpos, M., Gaganis, C., & Pasiouras, F. (2005). Explaining qualifications in audit reports using a support vector machine methodology. *Intelligent Systems in Accounting, Finance & Management, 13*(4), 197-215.

Dowling, C., & Leech, S. (2007). Audit support systems and decision aids: Current practice and opportunities for future research. *International Journal of Accounting Information Systems, 8*, 92-116.

Druică, E., Oancea, B., & Vâlsan, C. (2018). Benford's law and the limits of digit analysis. *International Journal of Accounting Information Systems, 31*, 75-82.

Dutta, I., Dutta, S., & Raahemi, B. (2017). Detecting financial restatements using data mining techniques. *Expert Systems with Applications, 90*, 374-393.

Dzuranin, A. C., & Mălăescu, I. (2016). The current state and future direction of IT audit: Challenges and opportunities. *Journal of Information Systems, 30*(1), 7-20.

Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. *Business Horizons, 58*(5), 493-500.

Eberle, R. (2017). Die Zukunft der Wirtschaftsprüfung. *EXPERTFOCUS, 11*, 751-755.

Eining, M. M., Jones, D. R., & Loebbecke, J. K. (1997). Reliance on decision aids: An examination of auditors' assessment of management fraud. *Auditing: A Journal of Practice & Theory, 16*(2), 1-19.

Eriksson, P., & Kovalainen, A. (2016). *Qualitative methods in business research.* (2nd ed.). London, United Kingdom: Sage.

Ernst & Young. (2018). How artificial intelligence will transform the audit. Retrieved December 28, 2018, from https://www.ey.com/en_gl/assurance/how-artificial-intelligence-will-transform-the-audit

Ernst & Young. (2019). How the EY Digital Audit is evolving to meet new challenges. Retrieved August 28, 2019, from https://www.ey.com/en_gl/audit/technology/helix

EXPERTsuisse. (2015). *Schweizer Handbuch der Wirtschaftsprüfung*. Zurich: EXPERTsuisse.

FAOA. (2017). Annual Report 2016. Retrieved June 30, 2019, from https://www.rab-asr.ch/#/page/127

FAOA. (2018). Annual Report 2017. Retrieved June 30, 2019, from https://www.rab-asr.ch/#/page/127

FAOA. (2019a). Annual Report 2018. Retrieved June 30, 2019, from https://www.rab-asr.ch/#/page/127

FAOA. (2019b). Website. Retrieved June 30, 2019, from https://www.rab-asr.ch/#/

Federal Statistical Office. (2018). Labour costs. Retrieved May 28, 2019, from https://www.bfs.admin.ch/bfs/en/home/statistics/work-income/wages-income-employment-labour-costs/labour-costs.html

Finlay, S. (2017). *Artificial intelligence and machine learning for business. a no-nonsense guide to data driven technologies.* (2nd ed.). London, United Kingdom: Relativistic.

Fontana, M., Birrer, T., Hummel, P., & Füchslin, C. (2018). Prüfungshonorare aus Sicht der Geprüften - Dritte "Audit Cost Survey". *Expert Focus, 4*, 257-263.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254-280.

Gaganis, C. (2009). Classification techniques for the identification of falsified financial statements: a comparative analysis. *Intelligent Systems in Accounting, Finance & Management, 16*(3), 207-229.

Gaganis, C., Pasiouras, F., Spathis, C., & Zopounidis, C. (2007). A comparison of nearest neighbours, discriminant and logit models for auditing decisions. *Intelligent Systems in Accounting, Finance & Management, 15*(1/2), 23-40.

Gartner. (2019). Gartner Hype Cycle. Retrieved November 10, 2019, from https://www.gartner.com/en/research/methodologies/gartner-hype-cycle

Gee, S. (2014). *Fraud and fraud detection: a data analytics approach*. New York, NY: John Wiley & Sons.

Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting, 29*(6), 536-555.

Gepp, A., O'Neill, T. J., Linnenluecke, M. K., & Smith, T. (2018). Big data techniques in auditing research and practice: Current trends and future opportunities. *Journal of Accounting Literature, 40*, 102-115.

Ghosh, S., & Gunning, D. (2019). *Natural language processing fundamentals*. Birmingham, United Kingdom: Packt Publishing.

Gibson, W. (1999). The science in science fiction. Retrieved November 12, 2019, from https://www.npr.org/2018/10/22/1067220/the-science-in-science-fiction

Gläser, J., & Laudel, G. (2009). Wenn zwei das Gleiche sagen... Qualitätsunterschiede zwischen Experten. In A. Bogner, B. Littig, & W. Menz (Eds.), *Experten-interviews, Theorien, Methoden, Anwendungsfelder* (3rd ed., pp. 137-158). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Gläser, J., & Laudel, G. (2010). *Experteninterviews und qualitative Inhaltsanalyse. als Instrumente rekonstruierender Untersuchungen. Jochen Gläser ; Grit Laudel* (4th ed.). Wiesbaden, Germany: VS Verlag für Sozialwissenschaften.

Goel, S., & Gangolly, J. (2012). Beyond the numbers: Mining the annual reports for hidden cues indicative of financial statement fraud. *Intelligent Systems in Accounting, Finance & Management, 19*(2), 75-89.

Goldshteyn, M., Gabriel, A., & Thelen, S. (2013). *Massendatenanalysen in der Jahresabschlussprüfung. Grundlagen und praktische Anwendungen mit Hilfe von IDEA*. Düsseldorf, Germany: IDW Verlag.

Günther, C. W., & Van der Aalst, W. (2007). *Fuzzy mining - adaptive process simplification based on multi-perspective metrics.* Paper presented at the 5th International Conference on Business Process Management (BPM 2007), Brisbane, Australia.

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review, 61*(4), 5-14.

Hail, L., Tahoun, A., & Wang, C. (2018). Corporate scandals and regulation. *Journal of Accounting Research, 56*(2), 617-671.

Hair, J. F. (2011). *Essentials of business research methods.* (2nd ed.). Armonk, NY: Sharpe.

Han, J., Kamber, M., & Pei, J. (2012). *Data mining. concepts and techniques.* (3rd ed.). Burlington, VT: Elsevier.

Hardeniya, N., Perkins, J., Chopra, D., Joshi, N., & Mathur, I. (2016). *Natural language processing: Python and NLTK*. Birmingham, United Kingdom: Packt Publishing.

Harvard Business Review. (2012). When you can't innovate, copy. Retrieved November 10, 2019, from https://hbr.org/2012/05/when-you-cant-innovate-copy

Hayes, R. S. (2005). *Principles of auditing: An introduction to international standards on auditing* (2nd ed.). Essex, United Kingdom: FT Prentice-Hall.

Helfert, M. (2000). Massnahmen und Konzepte zur Sicherung der Datenqualität. In R. Jung & R. Winter (Eds.), *Data Warehousing Strategie* (pp. 61-77). Berlin, Germany: Springer.

Hoogs, B., Kiehl, T., Lacomb, C., & Senturk, D. (2007). A genetic algorithm approach to detecting temporal patterns indicative of financial statement fraud. *Intelligent Systems in Accounting, Finance & Management, 15*(1/2), 41-56.

Huang, S.-Y., Tsaih, R.-H., & Lin, W.-Y. (2014). Feature extraction of fraudulent financial reporting through unsupervised neural networks. *Neural Network World, 5*, 539-560.

Huerta, E., & Jensen, S. (2017). An accounting information systems perspective on data analytics and big data. *Journal of Information Systems, 31*(3), 101-114.

Humpherys, S. L., Moffitt, K. C., Burns, M. B., Burgoon, J. K., & Felix, W. F. (2011). Identification of fraudulent financial statements using linguistic credibility analysis. *Decision Support Systems, 50*(3), 585-594.

IAASB. (2016). Exploring the growing use of technology in the audit, with a focus on data analytics. Retrieved July 30, 2019, from https://www.ifac.org/system/files/publications/files/IAASB-Data-Analytics-WG-Publication-Aug-25-2016-for-comms-9.1.16.pdf

IAASB. (2019a). *International Standards on Auditing*. New York, NY: IFAC.

IAASB. (2019b). Technology and the future ready auditor. Retrieved August, 29 2019, from https://www.ifac.org/system/files/publications/files/IAASB-Tech-Talk-May-2019.pdf

International Accounting Bulletin. (2017). Benchmarking intelligence for the accounting industry. Retrieved January 7, 2019, from http://www.crowe.ie/wp-content/uploads/2017/02/IAB-2017_World-Survey.pdf

Issa, H., Ting, S., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting, 13*(2), 1-20.

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*, 1-20.

Jans, M., Alles, M. G., & Vasarhelyi, M. A. (2014). A field study on the use of process mining of event logs as an analytical procedure in auditing. *The Accounting Review, 14*, 1751-1773.

Jans, M., Lybaert, N., & Vanhoof, K. (2010). Internal fraud risk reduction: Results of a data mining case study. *International Journal of Accounting Information Systems, 11*(1), 17-41.

Jans, M., van der Werf, J. M., Lybaert, N., & Vanhoof, K. (2011). A business process mining application for internal transaction fraud mitigation. *Expert Systems with Applications, 38*(10), 13351-13359.

Jiang, Y., & Jones, S. (2018). Corporate distress prediction in China: a machine learning approach. *Accounting and Finance, 58*, 1063-1109.

Kantardzic, M. (2011). *Data mining. concepts, models, methods, and algorithms* (2nd ed.). New York, NY: John Wiley & Sons.

Kartscher, P., Rossi, B., & Suter, D. (2013). *Wirtschaftsprüfung - interne und externe Revision. Systematischer Überblick für Wirtschaftsprüfer, Verwaltungsrat und Geschäftsleitung*. Zürich, Switzerland: Schulthess.

Katsis, C. D., Goletsis, Y., Boufounou, P. V., Stylios, G., & Koumanakos, E. (2012). Using ants to detect fraudulent financial statements. *Journal of Applied Finance and Banking, 2*(6), 73-81.

Kelleher, J. D., & Tierney, B. (2018). *Data science*. Cambridge, MA: The MIT Press.

Kiang, M. Y. (2001). Extending the Kohonen self-organizing map networks for clustering analysis. *Computational Statistics & Data Analysis, 38*(2), 161-180.

King, N., Cassell, C., & Symon, G. (1994). *Qualitative methods in organizational research: A practical guide*. London, United Kingdom: Sage Publications Ltd.

Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007a). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications, 32*(4), 995-1003.

Kirkos, E., Spathis, C., Nanopoulos, A., & Manolopoulos, Y. (2007b). Identifying qualified auditors' opinions: A data mining approach. *Journal of Emerging Technologies in Accounting, 4*, 183-197.

Klauser, M. (2019). Eine Branche in Bewegung: EXPERTsuisse Digitalisierungsstudie 2019 – Transformation von Datenverarbeitern hin zu Datenanalysten. Retrieved November 4, 2019, from https://www.expertsuisse.ch/files/cfdocument/news_3023.pdf

Klauser, M., & Herzog, D. (2017). Digitalisierung der Treuhandbranche. *EXPERTFOCUS, 11*, 832-837.

Knechel, W. R. (2007). *Auditing, assurance & risk* (3rd ed.). Mason, OH: Thomson / South-Western.

Knopf, J. W. (2006). Doing a literature review. *PS: Political Science & Politics, 39*(1), 127-132.

Knorr, E. M., & Ng, R. T. (1999). Finding intensional knowledge of distance-based outliers. *VLDB.* Retrieved September 18, 2019, from http://vldb.org/conf/1999/P21.pdf

Knox, S. W. (2018). *Machine learning*. New York, NY: John Wiley & Sons, Inc.

Koh, H. C., & Low, C. K. (2004). Going concern prediction using data mining techniques. *Managerial Auditing Journal, 19*(3), 462-476.

Koskivaara, E. (2000). Artificial neural network models for predicting patterns in auditing monthly balances. *The Journal of the Operational Research Society, 51*(9), 1060-1069.

Koskivaara, E., & Back, B. (2007). Artificial neural network assistant (ANNA) for continuous auditing and monitoring of financial data. *Journal of Emerging Technologies in Accounting, 4*, 29-45.

Kotsiantis, S., & Kanellopoulos, D. (2008). *Multi-instance learning for predicting fraudulent financial statements*. Paper presented at the 3rd 2008 International Conference on Covergence and Hybrid Information Technology. Conference Paper retrieved from https://ieeexplore.ieee.org/abstract/document/4682067

Kotsiantis, S., Koumanakos, E., Tzelepis, D., & Tampakas, V. (2006a). Financial application of neural networks: Two case studies in greece. In S. Kollias, A. Stafylopatis, W. Duch, & E. Oja (Eds.), *Artificial Neural Networks - ICANN 2006* (pp. 672-681). Athens, Greece.

Kotsiantis, S., Koumanakos, E., Tzelepis, D., & Tampakas, V. (2006b). *Predicting fraudulent financial statements with machine learning techniques*. Paper presented at the Advances in Artificial Intelligence: 4th Helenic Conference on AI, SETN 2006, Proceedings. Conference Paper retrieved from https://link.springer.com/chapter/10.1007/11752912_63

Kotu, V., & Deshpande, B. (2014). *Predictive analytics and data mining : Concepts and practice with RapidMiner*. San Francisco, CA: Elsevier Science & Technology.

KPMG. (2017). Clarity on Dynamic Audit. Retrieved September 15, 2019, from https://assets.kpmg/content/dam/kpmg/ch/pdf/clarity-on-dynamic-audit-en.pdf

Krieger, F., & Drews, P. (2018). Leveraging Big Data and Analytics for Auditing: Towards a Taxonomy. from https://www.researchgate.net/profile/Paul_Drews/publication/328902212_Leveraging_Big_Data_and_Analytics_for_Auditing_Towards_a_Taxonomy/links/5beaa034a6fdcc3a8dd2122d/Leveraging-Big-Data-and-Analytics-for-Auditing-Towards-a-Taxonomy.pdf

Küng, D., Hamberger, B., & Schweitzer, R. (2017). Datenanalysen in der Abschlussprüfung. *EXPERTFOCUS, 11*, 773-778.

Lantz, B. (2015). *Machine Learning with R* (2nd ed.). Birmingham, United Kingdom: Packt Publishing.

Larose, D. T., & Larose, C. D. (2015). *Data mining and predictive analytics.* (2nd ed.). New York, NY: John Wiley & Sons.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Leimeister, J. M. (2015). *Einführung in die Wirtschaftsinformatik* (12th ed.). Berlin, Germany: Springer.

Lenard, M. J., Alrm, P., & Booth, D. (2000). An Analysis of Fuzzy Clustering and a Hybrid Model for the Auditor's Going Concern Assessment. *Decision Sciences, 31*(4), 861-884.

Lindner, M. (2019). Die Diagnose kommt vom Computer. *NZZ.* Retrieved October 28, 2019, from https://www.nzz.ch/wissenschaft/ki-in-der-medizin-hilfe-bei-einfachen-und-repetitiven-aufgaben-ld.1497525

Liu, Q. (2014). *The application of exploratory data analysis in auditing.* Rutgers University-Graduate School-Newark, Newark, NJ.

Louwers, T. J., Blay, A. D., Sinason, D. H., Strawser, J. R., & Thibodeau, J. C. (2018). *Auditing & assurance services* (7th ed.). New York, NY: McGraw-Hill Education.

Makhzani, A., Shlens, J., Jaitly, N., Goodfellow, I., & Frey, B. (2015). Adversarial autoencoders. Retrieved September 20, 2019, from https://arxiv.org/pdf/1511.05644.pdf

Mann, P. S. (2001). *Introductory statistics* (4th ed.). New York, NY: Wiley.

Marten, K.-U., Quick, R., & Ruhnke, K. (2015). *Wirtschaftsprüfung. Grundlagen des betriebswirtschaftlichen Prüfungswesens nach nationalen und internationalen Normen* (5th ed.). Stuttgart, Germany: Schäffer-Poeschel Verlag.

Martens, D., Baesens, B., Vanthienen, J., Bruynseels, L., & Willekens, M. (2008). Predicting going concern opinion with data mining. *Decision Support Systems, 45*(4), 765-777.

Mautone, G., Häni, R., & Völker-Alber, J.-H. (2018). Der Schein (be-)trügt. Retrieved June 25, 2019, from https://www.pwc.ch/de/press-room/press-releases/pwc_mm_gecs2018_de.pdf

Mendes, A., Cardoso, R. L., Mário, P. C., Martinez, A. L., & Ferreira, F. R. (2014). Insolvency prediction in the presence of data inconsistencies. *Intelligent Systems in Accounting, Finance & Management, 21*(3), 155-167.

Mertens, P., Bodendorf, F., König, W., Schumann, M., Hess, T., & Buxmann, P. (2017). *Grundzüge der Wirtschaftsinformatik* (12th ed.). Berlin, Germany: Springer.

Meuser, M., & Nagel, U. (1991). ExpertInneninterviews - vielfach erprobt, wenig bedacht : ein Beitrag zur qualitativen Methodendiskussion. In D. Garz & K. Kraimer (Eds.), *Qulitativ-Empirische Sozialforschung Konzepte, Methoden, Analysen* (pp. 441-471). Opladen, Germany: Westdeutscher Verlag.

Meuser, M., & Nagel, U. (2002). ExpertInneninterviews - vielfach erprobt, wenig bedacht ein Beitrag zur qualitativen Methodendiskussion. In A. Bogner, B. Littig, & W. Menz (Eds.), *Das Experteninterview : Theorie, Methode, Anwendung* (pp. 71-93). Berlin, Germany: Springer.

Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In A. Bogner, B. Littig, & W. Menz (Eds.), *Interviewing experts* (pp. 17-42). Basingstoke, United Kingdom: Palgrave Macmillan.

Meyer, J. (2012). *Forensische Datenanalyse. Dolose Handlungen im Unternehmen erkennen und aufdecken*. Berlin, Germany: Erich Schmidt.

Moeller, R. R. (2011). *COSO enterprise risk management: Establishing effective governance, risk, and compliance processes* (2nd ed.). New York, NY: John Wiley.

Moeller, R. R. (2013). *Executive's guide to COSO internal controls: Understanding and implementing the new framework*. New York, NY: John Wiley.

Moore, G. A., & McKenna, R. (1999). *Crossing the chasm*. New York, NY: HarperCollins.

Moreira, J., De Carvalho, A., & Horváth, T. (2019). *A general introduction to data analytics.* New York, NY: John Wiley & Sons, Inc.

Nigrini, M. J. (1999). I've got your number. *Journal of accountancy, 187*(5), 79-83.

Nigrini, M. J., & Miller, S. J. (2009). Data diagnostics using second-order tests of Benford's law. *Auditing: A Journal of Practice & Theory, 28*(2), 305-324.

Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing, 16*(2), 52.

Omoteso, K. (2012). The application of artificial intelligence in auditing: Looking back to the future. *Expert Systems with Applications, 39*, 8490-8495.

Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory, 30*(2), 19-50.

Perols, J. L., Bowen, R. M., Zimmermann, C., & Samba, B. (2017). Finding needles in a haystack: Using data analytics to improve fraud prediction. *Accounting Review, 92*(2), 221-245.

Porter, B., Simon, J., & Hatherly, D. (2011). *Principles of external auditing* (3rd ed.). New York, NY: Wiley.

Provost, F., & Fawcett, T. (2013). *Data science for business* (1st ed.). Sebastopol, CA: O'Reilly.

Purda, L., & Skillicorn, D. (2015). Accounting variables, deception, and a bag of words: Assessing the tools of fraud detection. *Contemporary Accounting Research, 32*(3), 1193-1223.

PwC. (2012). Halo for Journals. Retrieved August 19, 2019, from http://halo.pwc.com/

PwC. (2019). Harnessing the power of AI to transform the detection of fraud and error. Retrieved September 15, 2019, from https://www.pwc.com/gx/en/about/stories-from-across-the-world/harnessing-the-power-of-ai-to-transform-the-detection-of-fraud-and-error.html

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers* (2nd ed.). Oxford, United Kingdom: UK Blackwell.

Roiger, R. J. (2017). *Data mining: A tutorial-based primer* (2nd ed.). Boca Raton, FL: CRC Press.

Rubin, H. J., & Rubin, I. S. (2012). *Qualitative interviewing. the art of hearing data.* (3rd ed.). Thousand Oaks, CA: Sage.

Ruhnke, K. (2019). Auf dem Weg zu Big Data Analytics in der Abschlussprüfung: Auswirkungen und Herausforderungen. *WPg, 72*(2), 64-71.

Runkler, T. A. (2016). *Data Analytics: Models and algorithms for intelligent data analysis* (2nd ed.). Wiesbaden, Germany: Springer Fachmedien.

Salijeni, G., Samsonova-Taddei, A., & Turley, S. (2019). Big data and changes in audit technology: Contemplating a research agenda. *Accounting & Business Research, 49*(1), 95-119.

Sauer, J., & Schumann, R. (2007). Modelling and solving workforce scheduling problems. Retrieved October 12, 2019, from https://www.researchgate.net/profile/Rene_Schumann/publication/200022584_Modelling_and_Solving_Workforce_Scheduling_Problems/links/0f31753b1c7b3a9b56000000/Modelling-and-Solving-Workforce-Scheduling-Problems.pdf

Schneider, F., & Bradtke, S. (2017). Digitalisierung: Auswirkungen auf die Revisionsaufsicht. *Expert Focus, 11*, 761-765.

Schneider, G., Jun, D., Janvrin, D., Ajayi, K., & Raschke, R. (2015). Infer, predict, and assure: Accounting opportunities in data analytics. *Accounting Horizons, 29*(3), 719-742.

Schreyer, M., Sattarov, T., Borth, D., Dengel, A., & Reimer, B. (2017). Detection of anomalies in large scale accounting data using deep autoencoder networks. Retrieved July 15, 2019, from https://arxiv.org/pdf/1709.05254.pdf

Schreyer, M., Sattarov, T., Borth, D., Dengel, A., & Reimer, B. (2018). Künstliche Intelligenz in der Wirtschaftprüfung: Identifikation ungewöhlicher Buchungen in der Finanzbuchhaltung. *WPg, 71*(11), 674-680.

Schreyer, M., Sattarov, T., Reimer, B., & Borth, D. (2019a). Adversarial learning of deepfakes in accounting. Retrieved November 2, 2019, from https://www.alexandria.unisg.ch/258090/1/NeurIPS_2019_RAIFS_final.pdf

Schreyer, M., Sattarov, T., Schulze, C., Reimer, B., & Borth, D. (2019b). Detection of accounting anomalies in the latent space using adversarial autoencoder neural networks. Retrieved August 15, 2019, from https://arxiv.org/pdf/1908.00734.pdf

Schumpeter, J. (1942). *Capitalism, socialism and democracy*. Stuttgart, Germany: UTB.

Schwab, K. (2017). *The fourth industrial revolution*. London, United Kingdom: Penguin Books.

Shafeeq, A., & Hareesha, K. (2012). *Dynamic clustering of data with modified k-means algorithm.* Paper presented at the Proceedings of the 2012 Conference on Information and Computer Networks.

Singh, N., Lai, K. h., Vejvar, M., & Cheng, T. C. E. (2019). Data-driven auditing: A predictive modeling approach to fraud detection and classification. *Journal of Corporate Accounting & Finance, 30*(3), 64.

Song, X. P., Hu, Z. H., Du, J. G., & Sheng, Z. H. (2014). Application of machine learning methods to risk assessment of financial statement fraud: Evidence from China. *Journal of Forecasting, 33*(8), 611-626.

Sonnenberg, C., & Vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Berlin, Germany: Springer.

Suh, I. S., & Headrick, T. C. (2010). A comparative analysis of the bootstrap versus traditional statistical procedures applied to digital analysis based on Benford’s law. *Journal of Forensic and Investigative Accounting, 2*(2), 144-175.

Tackett, J. A. (2013). Association rules for fraud detection. *Journal of Corporate Accounting & Finance, 24*(4), 15-22.

Thanaki, J. (2017). *Python natural language processing. explore NLP with machine learning and deep learning techniques*. Birmingham, United Kingdom: Packt Publishing.

Thiprungsri, S., & Vasarhelyi, M. A. (2011). Cluster analysis for anomaly detection in accounting data : An audit approach. *The International Journal of Digital Accounting Research, 11*, 69-84.

Titera, W. R. (2013). Updating audit standard - enabling audit data analysis. *Journal of Information Systems, 27*(1), 325-331.

Torgo, L. (2007). Resource-bounded fraud detection. In J. Neves, M. F. Santos, & J. M. Machado (Eds.), *Progress in artificial intelligence* (pp. 449-460). Berlin, Germany: Springer.

Torgo, L. (2011). *Data mining with R: Learning with case studies.* Boca Raton, FL: CRC Press.

Tremblay, M. C., Hevner, A. R., & Berndt, D. J. (2010). The use of focus groups in design science research. In A. Hevner & S. Chatterjee (Eds.), *Design research in information systems* (pp. 121-143). Berlin, Germany: Springer.

Vaisman, A., & Zimányi, E. (2014). *Data warehouse systems: Design and implementation*. Berlin, Germany: Springer.

Van der Aalst, W. (2011). *Process mining. discovery, conformance and enhancement of business processes*. Berlin, Germany: Springer.

Van der Aalst, W. (2016). *Process mining: Data science in action* (2nd ed.). Berlin, Germany: Springer.

Van der Aalst, W., & Weijters, A. J. M. M. (2005). Process mining. In M. Dumas, W. van der Aalst, & A. Ter Hofstede (Eds.), *Process aware information systems: Bridging people and software through process technology*. New York, NY: Wiley.

Vandervelde, S. D. (2006). The importance of account relations when responding to interim audit testing results. *Contemporary Accounting Research, 23*(3), 789-821.

Wang, R. W., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5-33.

Wang, T., & Cuthbertson, R. (2014). Eight issues on audit data analytics we would like researched. *Journal of Information Systems, 29*(1), 155-162.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association, 58*(301), 236-244.

Weiss, G. (2004). Mining with rarity: A unifying framework. Retrieved August 28, 2019, from http://delivery.acm.org/10.1145/1010000/1007734/p7-weiss.pdf?ip=130.82.70.29&id=1007734&acc=ACTIVE%20SERVICE&key=FC66C24E42F07228%2E191764A3709CF368%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1574895261_d195123caa44986d2dba676077cf0d89

Weissenberger, B., Förster, G., Bravidor, M., & Wesser, M. (2019). Wohin führt die Digitalisierung? *WPg, 20*, 1118-1124.

Werner, M., & Gehrke, N. (2015). Multilevel process mining for financial audits. *IEEE Transactions on Services Computing, 8*(6), 820-832.

Westermann, C., & Spindler, C. (2017). Anomalie-Erkennung in der Wirtschaftsprüfung mithilfe von Machine Learning. *EXPERTFOCUS, 11*, 791-797.

Whiting, D. G., Hansen, J. V., McDonald, J. B., Albrecht, C., & Albrecht, W. S. (2012). Machine learning methods for detecting patterns of management fraud. *Computational Intelligence, 28*(4), 505-527.

Wilson, J. S. (2009). *Essentials of business research: A guide to doing your research project*. London, United Kingdom: Sage.

Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a standard process model for data mining*. Paper presented at the Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining.

World Economic Forum. (2015). Deep shift: Technology tipping points and societal impact. Retrieved December 21, 2018, from http://www3.weforum.org/docs/WEF_GAC15_Technological_Tipping_Points_report_2015.pdf

World Economic Forum. (2019). Switzerland has the most highly skilled workers in the world. This is why. Retrieved October 28, 2019, from https://www.weforum.org/agenda/2019/10/switzerland-economy-skills-workforce-competitive/

Wu, J., & Coggeshall, S. (2012). *Foundations of predictive analytics*. London, United Kingdom: CRC Press LLC.

Yin, R. K. (2009). *Case study research: Design and methods.* (4th ed. Vol. 5). Thousand Oaks, CA: Sage.

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2018). Process-mining-enabled audit of information systems: Methodology and an application. *Expert Systems with Applications, 110*, 80-92.

Appendix

List of reviewed journals

- Abacus
- Accounting and Business Research
- Accounting and Finance
- Accounting and the Public Interest
- Accounting Forum
- Accounting Horizons
- Accounting in Europe
- Accounting Perspectives
- Accounting Research Journal
- Accounting Review
- Accounting, Auditing, & Accountability Journal
- Advances in Accounting
- Advances in Management Accounting
- Auditing: A Journal of Practice & Theory
- Australian Accounting Review
- Contemporary Accounting Research
- Controller Magazin
- Controlling & Management Review (before: Zeitschrift für Controlling und Management ZfCM)
- Der Betrieb
- Die Wirtschaftsprüfung (WPg)
- European Accounting Review
- EXPERTFOCUS (before: Schweizer Treuhänder)
- Financial Accountability and Management

- Intelligent Systems in Accounting, Finance & Management
- International Journal of Accounting
- International Journal of Accounting Information Systems
- International Journal of Accounting, Auditing and Performance Evaluation
- International Journal of Auditing
- International Journal of Managerial and Financial Accounting
- Journal of Accountancy
- Journal of Accounting Literature
- Journal of Accounting Research
- Journal of Accounting, Auditing & Finance
- Journal of Applied Accounting Research
- Journal of Business Finance & Accounting
- Journal of Corporate Accounting and Finance
- Journal of Forensic & Investigative Accounting
- Journal of Information Systems
- Journal of International Accounting Auditing and Taxation
- Journal of International Accounting Research
- Journal of International Financial Management and Accounting
- Journal of Management Accounting Research
- Management Accounting Quarterly
- Management Accounting Research
- Managerial Auditing Journal
- Research Journal of Finance and Accounting
- Review of Accounting and Finance
- Review of Accounting Studies
- The British Accounting Review
- Zeitschrift Interne Revision

Topic guide for audit companies

Segment	Questions
Assigned based on answer	What thoughts do you associate with data analytics in external auditing?
Methods and Tools	Within the auditing process, where do you use data analytics at the moment?
	Could you explain what kind of data analytics techniques are used?
Motivation	Why do you integrate data analytics into the course of an audit?
Application / Challenges	What experience have you had with data analytics so far, in particular, hurdles and successes?
	How do you ensure that the data analytics tools are adequately used or even used at all during an audit?
	How has the work process changed as a result of using data analytics in auditing?
	When or for what kind of customers is it worthwhile to use data analytics, and do you track your investments in data analytic tools?
	What example of future data analytic tools for auditing can you think of, in addition to the existing ones?

Audit Standard	To what extent does the current auditing standard influence the use of data analytics?
	Should the auditing standard be adapted and, if so, how?
	In literature it is often discussed that there is a shift from "sample" auditing to full audits with the help of data analytics. Nevertheless, it is still questionable how an auditor should deal with a large number of exceptional cases. For example, a sample of 40 transactions is audited and there are no exceptions, or perhaps a handful; if thousands or millions of transactions are checked, one could come across hundreds or thousands of exceptional cases. How should/can the profession deal with such a situation?

Topic guide for data analytics experts

Introductory question:

What are potential applications of data analytics for auditing?

Support question:

[If there are no possible uses for data analytics applications that come to the expert's mind] Regardless of auditing, what data analytic tools have you already created and for which area? Could you describe them?

In-depth questions for the suggested application:

- What preconditions must be met for this method, in particular with regard to the data? E.g., data quality; amount of data (especially in regards to training, testing and validation of the model); for supervised methods: assigning labels, data format.
- When creating a model one often has multiple choices with regard to attributes or datasets. How would you deal with this issue in the suggested application?
- What are the benefits of using this method for the auditor or his client?
- What are the limitations of this method?
- What hurdles could arise during the implementation of this application?
- How advanced does the user's technical knowledge have to be in order to apply this method or to interpret the results?
- How can the results of the method be transformed into a business context or into audit evidence, especially in a way that it is understandable for people without data science skills?

Curriculum vitae

Anita Stefanie Gierbl

Date of birth: October 3rd 1989 | Place of birth: Salzburg (Austria) |

Nationality: Austrian | E-Mail: anitagierbl@gmail.com

Professional experience

PwC Switzerland	Zurich, Switzerland
Assistant in Audit and Assurance	Since 12/2019
University of St. Gallen	St. Gallen, Switzerland
Research Associate	Since 04/2017
Swiss GAAP FER	St. Gallen, Switzerland
Technical Assistant	Since 04/2017
Bühler AG	Uzwil, Switzerland
Part time employee Internal Audit	09/2016 – 02/2017
Credit Suisse AG	Zurich, Switzerland
Internship Sales & Trading Services	07/2015 – 09/2015
PwC Austria	Vienna, Austria
Internship Capital Markets Accounting Advisory Services	06/2014 – 08/2014
Auxilia Accounting Inc.	New York, NY, USA
Part time employee Accounting and Taxation (remote)	01/2014 – 08/2014
Internship Accounting and Taxation	01/2013 – 12/2013
KPMG Austria	Linz, Austria
Part time employee Accounting Advisory Services	09/2011 – 12/2012
Internship Accounting Advisory Services	07/2011 – 09/2011

Education

University of St. Gallen	St. Gallen, Switzerland
Ph.D. in Management, Major Accounting	Since 09/2017
University of St. Gallen	St. Gallen, Switzerland
& **Melbourne Business School**	Melbourne, Australia
Master in Accounting and Finance	09/2014 – 05/2017
University of Applied Sciences Upper Austria	Steyr, Austria
& City University of Hong Kong	Hong Kong
Bachelor in Accounting, Controlling and Financial Management	09/2009 – 06/2012

Extracurricular activities at the University of St. Gallen

University of St. Gallen	St. Gallen, Switzerland
CAS in Teaching and Learning in Higher Education	Since 09/2018
DocNet – Doctoral Network	St. Gallen, Switzerland
President of the Executive Board	02/2018 – 03/2019
DocNet – Doctoral Network	St. Gallen, Switzerland
Member of the Executive Board for Finance	03/2017 – 02/2018