Improving the audit process by introducing analytical features to openESEA

Artur Moeijes
6904793
a.a.moeijes@students.uu.nl

First Supervisor: Dr. Sergio España Cubillo
Second Supervisor: Dr. Sietse Overbeek
Daily Supervisor: Vijanti Ramautar

A thesis presented for the degree of
Master of Science

Universiteit Utrecht

Utrecht University
The Netherlands
January 2022

Abstract

Worldwide, the duty for accountability is an important concern and thus sustainability reporting takes a central place in an organization. The information that is provided will also need to be audited. Not only does auditing have a positive effect on the organization, but it also ensures that the audited information is correct and reliable. This research contributes to explore what the audit and assurance process consists of on non-financial data, in other words ethical, social and environmental accounting (ESEA). To accomplish this, a generic model was developed through the examination of 10 different ESEA methods. The 10 methods were selected on specific criteria. Next, we looked at which activities in the process of the generic model could be improved. By analyzing the generic model, it was found that recommending specific information (indicators) for further audit can be an area for improvement. The improvements are performed by introducing analytical features, in particular Computer Assisted Audit Tools and Techniques. The CAATTs, were chosen based on the current state of the openESEA tool. Among various techniques, three techniques have been implemented in openESEA, namely scoring scheme, anomaly detection and visualization. Subsequently, we looked at the usability and applicability of the functionalities through internal and external validation with experts. All in all, this research has led towards improving the audit process and introducing appropriate analytical features in OpenESEA.

Keywords: Environmental reports, Auditing, Data mining, Anomalies, Analytics, Visualization, openESEA, Computer assisted auditing tools and technieues (CAATT)

Contents

1. Introduction 5
1.1. Observed problems . . . 6
1.2. Current situation . . . 6
1.3. Main goal . . . 7

2. Research method 8
2.1. Research Questions . . . 8
2.2. Problem investigation . . . 9
2.3. Treatment design . . . 9
2.4. Treatment validation . . . 9
2.5. Interview protocol . . . 10
2.6. Method Specification . . . 10
2.7. Method Comparison . . . 10
2.7.1. Method selection . . . 11
2.7.2. Method modelling . . . 11
2.7.3. Development of super-method . . . 11
2.7.4. Comparison of methods . . . 11

3. Background 12
3.1. ESEA . . . 13
3.1.1. Origin & Terminology . . . 13
3.1.2. Importance & Drivers . . . 14
3.1.3. Approaches / ESEA Methods . . . 16
3.1.4. Tools . . . 17
3.2. Audit . . . 17
3.2.1. Origin & Terminology . . . 17
3.2.2. Types . . . 20
3.2.3. Scope . . . 21
3.2.4. Importance & Drivers . . . 21
3.3. Assurance . . . 22
3.3.1. Levels . . . 22
3.3.2. Scope . . . 23
3.3.3. Importance & Drivers . . . 24
3.4. Computer-Assisted Audit Tools & Techniques . . . 25
3.4.1. Adoption . . . 26

4. State of the Art on Audit Method Fragments of Existing ESEA Methods 28
4.1. Method Selection . . . 28
4.2. Method Modelling . . . 30
4.3. Development of super-method . . . 31
4.4. Comparison of Methods . . . 31

4.5. Generic Model . . . 35

5. State of the art on Computer-Assisted Audit Tools and Techniques **37**
5.1. Benford's Law . . . 39
5.1.1. Mean Absolute Deviation . . . 41
5.1.2. Chi-Square Test . . . 42
5.2. Anomaly detection . . . 43
5.3. Visualization methods . . . 45
5.4. Scoring scheme . . . 46

6. Treatment Design **48**
6.1. OpenESEA . . . 48
6.2. Requirements and user stories . . . 49
6.3. Improving the fragments of the audit process by analytical features . . . 50
6.3.1. Concepts . . . 51
6.3.2. Indicator . . . 51
6.3.3. ESEA Method . . . 51

7. Treatment Validation **52**
7.1. Expert assessment . . . 52
7.2. Evaluation . . . 53

8. Discussion **54**
8.1. Limitations . . . 54
8.2. Future work . . . 55

9. Conclusion **56**

A. PDDs **69**
A.1. EMAS . . . 69
A.2. ISAE 3000 . . . 71
A.3. Green House Gases Protocol . . . 72
A.4. Natural Capital Protocol . . . 73
A.5. B Impact Assessment . . . 74
A.6. ISO 14011 . . . 75
A.7. AA1000 . . . 76
A.8. XES Social Balance audit . . . 77
A.9. FLA-SCI . . . 78
A.10.GRI . . . 79

B. PDD Validation Interview Protocol **80**

C. User Roles **84**

D. Screenshots of OpenESEA tool **89**

1. Introduction

Corporate Social Responsibility (CSR) and Environmental, Social and corporate Governance (ESG) are terms that has been used by businesses and academic researchers for decades [26, 145]. Corporate social responsibility is a manner of illustrating how companies measure and control their impact on the society. The increasing importance of having a good image or reputation to the society, stronger relationships with employees, and therefore better recruitment opportunities and a long-term sustainable business are key issues nowadays [3]. A comprehensive list of these motivations has been created by Ramautar [132].

Worldwide, the duty for accountability is an important concern. A study conducted in 2001 states that customers' expectation of a company's social and responsible behaviour is becoming increasingly important in the future [113]. It boosts employee satisfaction, increases creativity and innovation of employees and improves the public image. Companies must adhere to various agreements. As an example, there are agreements that before releasing sustainability reports, a company must have audited and assured the contents of the report [36]. But that's not the only motivation. We should not forget that the social aspect of a company is becoming increasingly important [127]. Thus, also the number of published sustainability reports is growing worldwide [154]. Some of these companies are independently assured by third parties and assures may or may not be from the auditing profession [138]. There are certain differences in the type of audit, the scope and the level of assurance. Chapter 3 will elaborate more on this in depth.

Sustainability reporting takes a central place in an organization. Additionally, it is important to note that for the publication of sustainability reports, different regulations and laws apply worldwide [74]. Currently, the European Parliament is negotiating to have companies officially include sustainability in their annual financial report by 2023 [36]. The proposal also introduces an audit requirement for the non-financial reports of these large companies. The main purpose for this audit requirement is ensuring that the reported information is complete and accurate [37].

Organizations take several new guidelines and regulations in their business strategy and improve their policy and accountability to the society on how the organization operates and performs in terms of sustainability. These organizations, especially multinationals, have to deal with many both international and local regulations. In particular, companies with a large number of employees, have to deal with sustainability reporting [1]. Aspects such as fair labor, diversity and inclusion are affected on a larger scale.

Auditing sustainability reports has positive effects. For instance, more transparency and trust regarding the performance of a company are created among the stakeholders. The topics addressed in the reports include aspects such as diversity, inclusion, human rights, minimum wage and co2 omission [4]. For stakeholders, publication of the reports and independent assurance means that the risk of incompleteness of data is significantly mitigated [91]

To bring these topics together, the term ethical social and environmental accounting (ESEA) will be used in this research. There are multiple methods used by organizations to perform ESEA. Examples of these methods are B Impact Assessment

(BIA), the Global Reporting Initiative (GRI), Common Good Balance Sheet, AA1000 and ISO26000. A relatively new tool, which will contribute to address such matters properly and bring these methods together, is the Ethical, Social and Environmental Accounting (ESEA) tool, also called openESEA, which is defined as 'the systematic, documented, periodic and objective evaluation of how well responsible consumers, enterprises, governments or countries are performing in terms of their impact in society and on the environmental responsibility' [55].

By using the existing infrastructures and related ESEA methods, regarding openESEA, an approach and structure to audit these reports is required. This approach, explores which audit fragments of the existing ESEA methods are suitable for auditing and assuring ESEA reports. The audit method fragment will be a improved functionality in openESEA to audit and assure ESEA reports. After the implementation, this research explores which analytical features, such as Digit Analysis, Machine Learning and Artificial Intelligence will be able to optimize the audit process. Nowadays, audit processes are supported by several Computer Aided Audit Tools and Techniques (CAATTs). CAATTs can be defined as 'any use of technology to assist in the completion of an audit' [124]. Advanced use of CAATTs can be seen as 'data analytics' and is increasingly being used across auditing profession because it saves time, improves the overall quality and saves unnecessary costs [5].

1.1. Observed problems

For the sake of clarity, the observed problems are explained in this section. Many methods for sustainability reporting exist, where some have implemented an audit process and others do not. Within openESEA tool there is not yet a auditing functionality designed nor implemented. The audit process could be designed as an (semi-)automated approach within the openESEA tool for conducting audits on sustainability reporting using different analytical approaches. In addition, the audit process could be extended with several analytical features that simplify the process for the auditor.

1.2. Current situation

Prior researchers have worked on openESEA and extended the tool. For instance, several additions have been made to the tool such as [96] researched and developed a feature to specify and deliver stakeholder surveys, [47] researched and developed a feature to specify and automatically generate infographics that present the account results and [132] layed all the groundwork for the development of an extendable openESEA tool. A beta version of the openESEA is constantly being improved and different functionalities to the tool are being added. However, the tool does not yet have a working audit and assurance function to verify the information generated by openESEA. This research will focus on adding the audit functionality and introducing analytical features such as predictive analytics, digit Analysis and machine learning to improve the audit process.

1.3. Main goal

The main goal of this project is to explore several ESEA methods and find the pain points in the audit process and improve these using the existent model-driven web-based ESEA tool. This research focuses on designing and implementing analytical functionalities in the openESEA tool. In order to improve the audit process, it is important to have an overview of all the included fragments of the process and step-by-step analyze the fragments with experts to find the pain points. The improvement should be carried out on the audit process fragment by researching possible existing CAATTs and applying them on openESEA tool. After the design and implementation of the audit functionality, an expert assessment will be set up to validate the implemented audit process. The validation of the added analytical features will be done by an expert assessment who will through the expertise in its field, review the work and additionally detect the inconsistencies.

2. Research method

This study aims to improve a specific fragment of an implemented audit process by decomposing into three research questions and sub questions, which can be found below. In this research project, the design cycle by Wieringa [157] will be followed. The design cycle consists of three stages: Problem Investigation, Treatment Design and Treatment Validation. Each stage contains one Research Question (RQ) and several sub questions. The activities, which are carried out during the research, and split into three stages of Wieringa, are shown in the Gantt Chart in the next chapter. The research was conducted in collaboration with one other Master of Business Informatics student, Ties van Dijk, who will be investigating the ESEA audit process as well, but who will be focusing on the managerial features of openESEA.

2.1. Research Questions

Problem investigation

RQ1: *What is the state of the art on audit method fragments of existing ESEA methods?*
RQ2: *What is the state of the art on analytical features in computer-assisted auditing tools and techniques (CAATTs)?*

Treatment design

RQ3: *How can the fragment of the audit process in openESEA be improved or (semi)-automated?*

Treatment validation

RQ4: *Validate the audit functionality of openESEA!*

RQ4.1: *How do experts assess the audit functionality of openESEA?*
RQ4.2: *How well does the analytical added functionality of openESEA perform?*

2.2. Problem investigation

To start at the beginning, *T01*, *T02* and *T03* are the activities which can be found in this document. *T01* is to *define the problem statement and motivation of the research* to conduct this research. As for now, the results of these activities can be found in section 1 as well as *the outline of the research method (T03).* Subsequently, in section 2 the *research questions are defined (T02).* Further, there will be an addition to the introduction where the literature review will be done on *researching the first state of art of audit processes (T04)*, in parallel *interviews* [61] *will be conducted (T05)* with experts in the field of auditing and *PDDs [148] will be created (T06)* based on the information obtained from both the interviews and literature review. The literature review is both scientific, grey and industrial literature. The second state of the art will be found by conducting a *literature review on Computer Aided Audit Tools and Techniques (CAATTs) (T07).* By engaging in this activity, knowledge and experience will be gained with the openESEA tool in the meantime. The literature review results in both state of the arts. At the end of the literature review, *preparation and presentation of the colloquium will be performed (T08 and T09)*

2.3. Treatment design

At the start of the Treatment Design, information gathered by conducting the interviews can be *converted into requirements (T10).* After the requirements are set up, they will be *validated and prioritized (T11 and T12)* by interviewing experts and further prioritized by using MoSCoW [2]. The requirements will be written in User Stories and based on these prioritized user stories, *openESEA analytical features for auditing will be developed (T13).* At the end of this stage, with the knowledge and experience gathered in the first stage, an (semi-) automated audit extension of the audit fragment(s) will be developed.

2.4. Treatment validation

During the last phase of this research, the results developed in the previous phases will be validated. This phase starts with *selecting and planning appropriate validation methods (T14).* After finding the best suitable validation methods, the *validation will be executed and determined in the paper (T15).* The most appropriate validation methods will be considered. Given the implementation of an audit process and its extension through analytical features, it is expected that an expert assessment and an experiment will be selected. Treatment validation, as well as the whole paper, will be resulting in the following activities (*T15-T20*), including *finalizing, preparing and presenting the final paper.*

2.5. Interview protocol

An interview protocol was developed in advance for conducting interviews. This protocol provides a structure that was applied during the interviews. The reason for using a protocol, is to bring some order to the interview and to avoid overlooking important or crucial elements. An interview protocol was used for all interviewees who had different functions. The goal of these interviews is to ensure that our interpretation of the documentation is correct and that we do not miss any essential activities or concepts. The purpose of these interviews is to provide confirmation that our representation of the documentation is accurate and that we are not overlooking essential activities or concepts. The structure of the interview can be found in Appendix B

The individuals interviewed had different functions, which are divided in:
- **Practitioner**, i.e. an employee of Big Four assurance firm
- **Policy maker or Method advisor**, that determines if the method needs to be revised, and applies adjustments if necessary. i.e. a party who provides input on the method or an advisor to the organisation that maintains the method.
- **Direct party or Method maintainer**, an employee directly involved in the development or expansion of the method.

2.6. Method Specification

First of all, the ESEA methods are specified by interpreting the information based on their documentation. The documentation is found during the literature review and is translated in PDDs. Some of the ESEA method documentation contains a specific audit method fragment in their own figures or reports. Appropriate ESEA methods that contain an audit fragment will be collected, according to the following predefined criteria:

1. existence of an audit component
2. recommendation by experts in the field

When additional documentation can be found on these audit fragments, this information is used and or replaced by uncertain activities. Six out of ten PDDs are validated during the interviews with the experts. Appendix A contains the PDDs including the author of each diagram. The diagrams and documentation are manually analysed and compared. These will result in a super-method.

2.7. Method Comparison

A method comparison is conducted to identify the main features and concepts of the audit method fragments of ESEA Methods. By conducting a method comparison, it will allow us to create insights into what fragments of the audit may be useful and important to support openESEA with. The method comparison is done by following the 4 steps based on Weerd et al [149]

2.7.1. Method selection

The first step to be performed is to collect useful and appropriate ESEA methods that contain an audit fragment or assurance fragment. These will be analyzed and compared at a later stage. The selection is based on discussions within the research group and both usability and relevance of the method. It is important to note that the method contains an audit fragment and that there is sufficient information about it in the documentation.

2.7.2. Method modelling

After the selection of the ESEA methods has taken place, we will go deeper into the documentation of the corresponding method. The reason behind this is to get a complete and clear picture of what the ESEA method exactly looks like. If the picture of the documentation is clear, Process Deliverable Diagrams of the selected methods will be developed or improved. The techniques of modeling will be performed according to the guidelines and rules of Van de Weerd and Brinkkemper [148].

2.7.3. Development of super-method

The following step in method comparison is the development of the super method. A super-method is defined as the smallest common denominator of activities and concepts in the meta-models [151]. This means, that a super method contains all the activities from all the methods that appear at least once. To build this method, it is important to oversee all the activities accuring in the methods and systematically include all the activities. If an activity seems to be almost the same as the one from another method, we discuss this together with the research group and come up with a decision.

2.7.4. Comparison of methods

After the super method is developed, we will compare the per ESEA method audit components to develop a generic model where the key elements of audit occur. An activity table will be developed with corresponding Super Method PDD. Then, based on the most common audit fragments, it will be condensed into Generic model. The activities are compared using the following notation, where 'a' is an activity from the super method and 'b' is an activity that is compared from an ESEA method [151].

- a '=' b: The activity/concept 'a' is equivalent to the activity/concept 'b'
- a '<' b: The activity 'a' does less than the activity 'b'
- a '>' b: The activity 'a' does more than the activity 'b'
- a '><' : A part of the activity 'a' overlaps a part of the activity 'b', and other parts do not overlap
- a '(empty)' b: The activity/concept 'a' is not equal/present to the activity/concept 'b'

3. Background

This chapter will elaborate on the background for this research. We will discuss four concepts in detail: ESEA, accounting, audit and assurance. From the outset, it is important to note that these concepts are related and that accounting, auditing and assurance are phases of the overarching ESEA process. The ESEA process starts with (1) accounting, followed by (2) audit, followed by (3) assurance [22]. This research focuses on the audit and assurance phases of the ESEA process. One should keep in mind, however, that companies may elect to skip the audit and assurance phases, and that these phases are thus not always executed [154].

ESEA is an acronym for 'ethical, social and environmental accounting'. As mentioned before, there is a plethora of terms that are used for the concept of ESEA, such as non-financial reporting, corporate social responsibility and more. In this research, the acronym ESEA is adopted, as this term both originates from scientific research and covers all ethical, social and environmental angles.

The exact structure of the ESEA process and whether auditing and assurance are part of it is the subject of some confusion, which is well documented in scientific literature [73]. At first glance, it may seem rather confusing that we consider the first phase of 'ethical, social and environmental accounting' to be accounting. However, there is a compelling argument for this approach, which is rooted in the methods that companies may use to perform ESEA.

Many ESEA methods prescribe an accounting, auditing and assurance phase [131]. This research aims to be of value to the sector as a whole, and thus, industry terminology should be taken into account. It is not the aim of this research to create a new definition for the overarching process. An initiative where researchers and industry come together to create new and universally usable terminology would be beneficial to the entire sector. However, this is outside the scope of this study. In short, while terminologically not strictly correct, we elected to structure the definitions by considering the ESEA process with three phases of accounting, audit and assurance. In the following paragraphs, we will elaborate on the exact definitions of these concepts.

To define the overarching ESEA process, we will adopt the definition of Espana et al: 'the process of assessing and reporting the ethical, social and environmental effects of a company's economic actions to particular interest groups with society and to society at large' [54]. This definition is based on that of Gray [73], but with the 'ethical' element added, to be able to cover the ethical governance aspect.

The first phase in the ESEA process is accounting. According to the Cambridge Dictionary, the definition of accounting in the financial sector is: 'the work of preparing the financial records of people, companies, or organizations' [48]. As this research concerns ethical social and environmental performance of companies, we will define accounting as: 'the process of preparing the non-financial reports of companies'. The result of the accounting phase is a non-financial report.

The next phase in the ESEA process is auditing. In this phase, the non-financial report resulting from the accounting phase is audited. According to the Cambridge Dictionary, the definition of audit is: 'an official examination of the accounts of a company' [49].

In an ethical, social and environmental context, we will define audit as: 'an official examination of the non-financial reports of a company'. The result of the audit is an auditor's report, which the Cambridge Dictionary defines as 'a formal document that states an auditor's judgment of a company's accounts' [50]. In an ethical, social and environmental context, we will define an auditor's report as: 'an auditor's judgment of a company's non-financial reports'.

The final phase in the ESEA process is assurance. The objective of assurance of non financial reports is to provide comfort to stakeholders about the accuracy of reported data [59]. Hence, the assurance provider *assures* that the reported data is accurate. Assurance is provided based on the auditor's report that is generated during auditing [22]. We will define assurance as: 'the process of assuring that the reported data is authentic and accurate'.

This chapter will be structured according to the overarching ESEA process. However, we will not discuss the first ESEA phase, accounting, in detail. The reasoning for this is that this research focuses specifically on the audit and assurance phases. Moreover, the accounting phase of the ESEA processes has already been extensively covered in previous research, for example by Ramautar [131], Derikx [47] and Kruiper [96]. We will also cover Computer-Assisted Audit Tools, which are used to support almost all auditing procedures involving data collection and analytics.

The origin, importance and approaches of ESEA are discussed in Section 3.1. Section 3.2 will discuss audit's importance, origin, types and scope. Section 3.3 will elaborate on different levels of assurance, different scopes and the importance to perform assurance. Finally, in Section 3.4, Computer-Assisted Audit Tools and Techniques and different CAATTs' adoptions are discussed.

3.1. ESEA

This section will elaborate on the origin and definition of ESEA, the drivers for an organisation to include ESEA in it's policy making decisions, ESEA approaches, and tools that support the ESEA process.

3.1.1. Origin & Terminology

While the demand and need for non-financial information has been steadily increasing in recent decades, the use and inclusion of social considerations and policies in organizational investment and strategy began to develop as early as 1898. One example of this is the Quaker Friends Fiduciary Corporation. For Quaker, these considerations included the avoidance of 'sin shares'. This amounted to a 'no guns, alcohol or tobacco' investment policy, designed to align their core business values with their investments [134]. It did not stop there, on the contrary, it developed more and nowadays, non-financial goals are increasingly high on the agenda for all kinds of organisations.

Businesses such as Ernst & Young (EY) and Deloitte highlight the growing importance of non-financial performance in regulation, supply chains and investment decisions [115, 6]. Socially Responsible Investment (SRI), 'ethical investing', 'social investing' and 'responsible investing' are terms that have been used by businesses and academic researchers to indicate a social investment strategy [27, 145]. Back in the 1950s, non financial information, then called Socially Responsible Investment, became a strategic position for investors and major corporations [65]. Over time, terms such as 'ethical investing', 'social investing' and 'responsible investing' have been replaced by SRI.

Since the work of the economist Howard Bowen in 1953, where he published a paper called 'Social Responsibilities of the Businessman', Bowen is being recognized as the founding father of Corporate Social Responsibility (CSR) [8, 20]. CSR has since become the term among companies and academics that embraces the idea that organizations bear responsibility not only to investors, but also to society. Thus, reporting and accountability to society also became an increasingly important aspect. Environmental, Social and corporate Governance (ESG) first appeared in a United Nations Global Compact report 'Who Cares Wins: Connecting Financial Markets to a Changing World' [114]. For this report, the former United Nations Secretary General invited a joint initiative of financial institutions "to create guidelines, regulations and recommendations on how to better integrate environmental, social and corporate governance issues in asset management, securities brokerage services and associated research functions" [51]. The topics addressed in the reports and accounting include aspects such as diversity, inclusion, human rights, minimum wage and co2 omission.

SRI, CSR and ESG are just a small selection of the terms that are used to address the same non-financial organizational performance. In practice these and other terms are used interchangeably. For example, Garriga still refers to CSR [69], Flower refers to 'Environmental Accounting' [75], and Gray refers to 'Integrated Reporting' [62]. For the sake of consistency in the terminology, in this research, the term Ethical Social and Environmental Accounting (ESEA) will be used. As mentioned before, ESEA is defined as the process of assessing and reporting the ethical, social and environmental effects of a company's economic actions to particular interest groups with society and to society at large [54].

3.1.2. Importance & Drivers

As briefly mentioned in the previous section, reporting and accountability to society has an important role within organizations. Organizations that partake in ESEA, hereinafter called 'reporting organizations', use non-financial reports to formalize their position in society and to assist their firms in developing good business practices [127]. ESEA has several other benefits to organisations. This section will discuss a non-exhaustive list of motivations for organisations to perform ESEA. A comprehensive list of these motivations has been created by Ramautar [131].

First of all, businesses might have a genuine interest on ethical, social and environmental aspects. In this case, businesses may perform ESEA to discover areas in which the business might improve [53]. However, it is important to note that the underlying motivation for discovering these areas of improvement might be extrinsic or intrinsic. In the case of extrinsic (or strategic) motivation, the business performs ESEA because this might 'contribute to financial success in the long run'. In the case of intrinsic (or moral) motivation, a business holds that ESEA 'is a moral duty of companies towards society' [72].

Regardless of the extrinsic or intrinsic nature of the motivation, companies might have already identified ethical, social and environmental areas of improvement. In this case, ESEA might be performed to discover the extent to which the organisation meets the associated ethical, social and environmental values. In a similar vein, ESEA can also be used to continuously track the performance of these values [98].

ESEA might not only be important for a company's performance, but also to its employees. The ethical, social and environmental performance of a company has an effect on its human capital. For example, studies have shown that reporting has an effect on employee satisfaction. In 2013, Kim et al. have shown that one of the main benefits of performing ESEA is an improvement in individual staff motivation [91]. Moreover, engaging in ESEA has a positive effect on retention intention of individual employees [99]. Lastly, multiple studies have suggested a relation between perceived ESEA performance and job applicant attraction [34, 133].

Apart from its effects on employees and company performance in general, ESEA also affects the public perception of the company [94, 82]. Perhaps most important is the effect of ESEA performance on customer perception. In 2019, Kim et al. have shown that ESEA performance has a positive effect on customers' knowledge, trust and perception on company reputation [93]. This relationship has been observed in multiple industries. Similar observations have been made in airlines [128], the hospitality industry [107], the banking sector [126] and other industries [101, 67, 31].

Taking into account the effect of ESEA performance on public perception, it stands to reason that companies would try to capitalize on its potential by employing it as a marketing apparatus. In 2001, leading research in the field suggested that marketing takes a lead role in ESEA related activities [97]. While a relation between customer perception and ESEA has been observed, as discussed in the previous paragraph, marketing regarding the subject has its complications. Especially the emergence of greenwashing, where companies mislead customers about their environmental performance or the environmental benefits of a product or service, has lead to consumer scepticism on ESEA marketing [46, 87].

One construct that is used in ESEA marketing efforts is that of third party certification [109]. Companies may want to obtain such third party certification to demonstrate their ESEA efforts [41]. Certification may also have other benefits. For example, in 1996, Creyer and Ross conducted a study where parents of middle-aged children were asked what the responsibility of a company should be in terms of ethical behavior, charity donations and impact on society. The results showed that customers are willing to pay more if the company takes its responsibility towards society and the environment into account in their business decisions [38]. Moreover, research suggests that customers are willing to pay more money for a product that has an ESEA-related certification [43, 140]. Some ESEA methods include assurance and certification, others make use of third-parties. Third-party certification may also come in the form of a network of responsible businesses of which the business wants to become (or is) a member of [11].

One example of such a network is the Economy for the Common Good[1].

Lastly, countries may oblige businesses to perform ESEA by means of legislation or other governmental regulation [82]. For example, in the European Union, Directive 2014/95/EU enforces large public interest entities to 'include non-financial statements as an integral part of their annual public reporting obligations' [36]. In April 2021 the Commission of the European Union proposed an amendment to Directive 2014/95/EU, which extends the scope of the directive [57]. In the amended directive, all large companies have to perform ESEA. A company is qualified as large when it conforms to at least two of the three following criteria:

- A balance sheet total of over €20 million;
- A net revenue of over €40 million;
- An average number of employees during the fiscal year of over 250.

Moreover, the proposal introduces an audit requirement for reported sustainability information, which will be discussed in Section 3.2.

3.1.3. Approaches / ESEA Methods

As Wallage argued back in 2000, non governmental organisations will play an increasing role in bringing structure to ESEA [154]. This statement has been validated in 2019, when a study found that a large number of NGOs have emerged with the goal of structuring the process and creating a framework for ESEA [131]. As a result, there are multiple methods by which companies can perform ESEA. Examples of these methods are the Global Reporting Initiative (GRI)[2], Common Good Balance Sheet [3], ISO14000[4], ISO26000[5] and the B Impact Assessment (BIA)[6]. Different methods have varying characteristics. For example, ISO14000 is focused on environmental performance, while ISO26000 emphasizes social performance.

Some methods enable businesses to obtain a certification [53]. For example, BIA enables companies to become a 'Certified B Corp'. Research has shown that some companies apply multiple ESEA methods [131]. One argument for companies to do so is to obtain multiple certifications. The certification for such methods is often issued by a monitoring organization, which is often the organisation that develops and maintains said method. We define a monitoring organisation as an organisation that develops and/or monitors initiatives, principles and/or standards related to corporate ethical, social and environmental performance [136]. Companies that apply the same method can also form a network of responsible organizations, which is often monitored by the monitoring organisation [131].

[1]https://www.ecogood.org/
[2]https://www.globalreporting.org
[3]https://www.ecogood.org
[4]https://www.iso.org/iso-14001-environmental-management.html
[5]https://www.iso.org/iso-26000-social-responsibility.html
[6]https://bimpactassessment.net/

3.1.4. Tools

Many ESEA methods are supported by information communication technology (ICT) tools. These tools are often developed by the monitoring organisation that developed the corresponding method, and thus support only a single method. Hence, when a company wants to apply multiple ESEA methods, they have to utilize multiple tools to do so. As methods are often based on the same data, this leads to redundancy in work [131]. This results in companies entering the same data in multiple tools. Utrecht University has developed the openESEA tool to combat this issue. openESEA is an easily extendable tool that supports multiple ESEA methods. The tool is developed in such a way that (monitoring) companies can easily add new ESEA methods to the tool [96].

Previously, researchers and master's students of Utrecht University have worked on openESEA and extended the tool. Several additions have been made. For example, Ramautar laid the groundwork for the development of the tool [131], Kruiper researched and developed features to specify and deliver stakeholder surveys [96], and Derikx researched and developed features to specify and automatically generate infographics that present the account results [47]. A beta version of openESEA is constantly being improved and different functionalities to the tool are being added. OpenESEA will be discussed in further detail in Section 6.1.

3.2. Audit

Audit is the second phase of the ESEA process. The result of the first phase, accounting, is the non-financial report. During the second phase, this non-financial report is audited. The result of the audit is an auditor's report, which is defined as a formal document that states an auditor's judgment of a company's non-financial reports. This section will discuss the origin and definition of auditing, it's place in modern society, the drivers for companies to perform auditing, and modern auditing approaches.

3.2.1. Origin & Terminology

The word 'audit' derives from the Latin word 'audire', which means 'to hear'. As mentioned in the introduction to this chapter, according to the Cambridge Dictionary, the contemporary definition of audit is: 'an official examination of the accounts of a business' [49]. The concept of auditing can be traced back to ancient civilizations in China, Egypt and Greece [21]. References to auditing can be found in works of Aristotle, but also in writings originating in medieval times. For example, in medieval Britain auditors heard the accounts read out and made sure that the organization's management did not act inappropriately or corrupt [108]. It is hypothesized that, before 1800, auditing was restricted to financial transactions, where every transaction was verified in detail [144].

In the 19th century, more large companies started to arise. However, the creation of these corporations called for a lot of financial capital. Much of this capital was provided by small individual investors, who were, due to legal obligations at the time, often liable on behalf of the company. These circumstances called for the protection of these investors. Hence, the financial audit process evolved to where it's main purpose was fraud detection. Moreover, the solvency of companies in the balance sheet was checked [130].

During the first half of the 20th century, the focus of the audit process shifted. This was caused by a change in economic circumstances. The economy, and thus the size of many companies grew rapidly. Moreover, company management and ownership became more and more separated, as the ownership of listed companies became more decentralized due to public issuing of shares. In these changing circumstances, the focus of auditing shifted towards providing credibility towards shareholders on the financial statements that were created by the management of a company [130]. This period also saw the advent of the use of sampling techniques in auditing, as well as evidence gathering through both internal and external sources [144].

The second half of the 20th century saw no great shifts in the focus of audit processes. There was, however, a growth in the reliance on sampling techniques. This was caused by the continuing growth of the world economy. This growth caused such a rise of transactions that sampling transaction for auditing became essential [135]. During this time period the concept of materiality was also introduced, which refers to the extend to which a transaction is important in a company's financial statements. Judgement on the materiality of a transaction influenced the decision to include the transaction in an audit [144].

Sampling procedure evolved further during the 1980s and 1990s, as risk-based auditing and analytical procedures were introduced. Risk-based auditing is an auditing approach where the auditor focuses on those transactions that contain the highest risk of error [147]. In this sense, the approach is related to materiality assessment, albeit with a focus on risk of error rather than the extend to which a transaction is important in a company's financial statements. Analytical procedures primarily involved identifying unexpected differences in financial statements. These might include account balance comparisons, trend analysis and comparisons of budget to actual [78].

From the 1990s to 2021, the auditing process evolved to where it is today. During this period, two important changes in the financial auditing process can be identified. First of all, nowadays, the auditor is expected to provide more value-added services next to verifying financial statements. Such services include business risk assessment and irregularity reporting [35]. The second, and arguably most important change has been the introduction of ICT tools that assist in the auditing process. Such Computer Assisted Auditing Tools and Techniques (CAATTs) might include analytical features that help in transaction sampling and transaction analysis, or managerial features that allow for managing the process as a whole [24].

So far, we have mostly covered financial audit, as this is the sector that the concept of auditing originates from. It is important to understand the context from which the concept originates. However, other forms of auditing have emerged over the years. Some examples are information technology audit [32], healthcare audit [85] and even innovation audit [30]. Most important, however, in the context of this study is the concept of ethical, social and environmental audit.

There are multiple developments that led to the concept of what we can now consider an ethical, social and environmental audit. Research has shown that separate ethical

[120], social [20] and environmental [106] audits can be identified. As an ethical, social and environmental audit covers all of these concepts, these individual audit processes have all been influential in the evolution of this audit.

Like the concept of ESEA, social audit can be traced back to the emergence of CSR. In his 1953 paper, Bowen coined the term 'social audit' [20]. However, at the time, social audit was mainly a tool for management to achieve insight in a company's social performance [28]. In other words, an auditor did not independently examine the non-financial reports of a company, but merely assessed social performance, similarly to ESEA in general. This resulted in an expectation gap concerning social audit, which ultimately led to a shift in audit focus. Similarly to financial audit, the aim of social audit became to improve the credibility and completeness in social reporting [9].

Environmental auditing developed similarly to social audit. The concept started of as a tool for management to achieve insight in a company's environmental performance [119]. However, a second purpose for environmental auditing evolved over time, with the aim to improve the credibility and completeness in environmental reporting [106]. Nowadays, environmental audits are used for both purposes, and the exact interpretation of the concept varies between companies and stakeholders [42].

In this research, we will adopt a definition of ethical, social and environmental auditing that focuses on providing credibility to a company's non-financial performance. The main argument for this is that the IT tools that this research aims to improve support the ESEA process, streamlining the process of non-financial reporting. Hence, auditing functionality is aimed at providing credibility to these financial reports.

We will define auditing as: 'an official examination of the non-financial reports of a company'. Moreover, from this point onwards, whenever we mention auditing, we imply ethical, social and environmental auditing, unless explicitly specified otherwise. One should note, however, that this is the definition of the container concept 'audit'. Even when restricting the context to an ethical, social and environmental one, an audit can have different *types* and *scopes*.

There are three types of audit: first party audit, second party audit and third party audit. The basic premise of these three types is that a first party audit is performed by the reporting organization itself, a second party audit is performed by a monitoring organization, and a third party audit is performed by an independent, third-party organization. Moreover, it is important to know, that an audit has a scope. In a ethical, social and environmental context, this can be a method and/or data scope. The different types and scopes of audit will be discussed in the next subsections.

3.2.2. Types

Audit can be distinguished and performed by three parties, first party, second party and third party. All three audits have different viewpoints and are performed by different actors. Figure 1 visually shows how the types of audit can be classified.

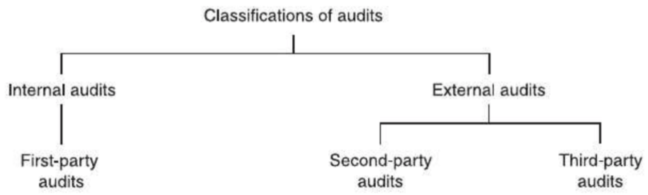

Figure 1: Classification of different audits, took from [95]

First party audit is performed by the reporting organisation itself. First party audit is also called internal audit. Internal audit serves as internal consultation. Moreover, there is no direct involvement in the preparation to process the results of the company's performance [112]. Independence with respect to management is an essential requirement of the internal audit. It is a requirement that ensures the internal auditor will be able to operate objectively and impartially without being submitted to the influence of the same authorities as the functions that is being audited [68].

Second party audit is carried out by one company to another and entails an audit undertaken by an entity with a trading or business relationship with the organisation [131]. Second party audit is a type of external audit. As described previously, many ESEA methods are managed and developed by a monitoring organization. For a company to get a certification or enter a group of sustainable organizations, the monitoring organization such as B corp or AA1000 itself might execute the audit. A number of reasons exist why an reporting organisation would want have a second party audit performed. Stephens and Roszak mention several examples such as provision of the input to selecting, grading and approving the reporting organization, a mechanism in the improvement of quality of their processes and results, and increasing mutual understanding of requirements [142]. The result of a second party audit may be an assurance statement, certification or improvement recommendation.

Third party audit is a type of audit which is being performed by an external, third party company. Prime examples of these companies are big accountancy firms such as Ernst & Young or Deloitte. These companies are free of any interests or conflicts and have no relationship in terms of provider-customer. Third party audit is a type of external audit. The outcome of the third party audits may be an assurance statement, recognition, a fine, a penalty, registration or certification by the third-party organization or an interested party [95].

3.2.3. Scope

Prior to the audit, an auditing organization needs to agree on a scope together with their customers. When auditing ethical, social and environmental performance, two different scopes can be identified: method and data audit.

Method audit examines whether an non-financial report or the information that is provided is consistent with the method that has been applied. For instance, AA1000 has three principles that are examined in-depth: inclusiveness, materiality and responsiveness. Inclusiveness considers whether the key stakeholders engage in the ESEA process. During materiality assessment, it is verified whether the company's report considers all significant ethical, social and environmental aspects and issues relevant to the organization. Lastly, during responsiveness assessment, it is examined whether the organization addresses the interests of all stakeholders and whether their concerns shape the organization's sustainability management and public relations [71].

Data audit includes an audit of all the metrics and information in a report. This includes performance indexes, statistical data and all other non-financial information. This is also by far the most common and widely used approach, and is the audit scope that people usually think of whenever auditing is mentioned [71]. Auditing data is done by retrieving evidence and comparing it to the data.

3.2.4. Importance & Drivers

Perego and Kolk have shown that the numbers of organisations having their sustainability reports audited increased from 21.4% in 1999 to 55.8% in 2008 [125]. The purpose of auditing is to convincingly demonstrate the reliability of the published report. Fundamentally, organizations must be held accountable for their actions, and this accountability must be verified through audit and assurance statements. As discussed in Section 3.1, one of the main motivations for companies to perform ESEA is to comply with legislation or other governmental regulation. Similarly, some jurisdictions oblige companies to have their non-financial reports audited. This is an audit driver that is both logical and very well documented in research [143].

In the European Union, Directive 2014/95/EU enforces large public interest entities to perform ESEA [36]. In April 2021, the European Commission proposed an amendment to Directive 2014/95/EU, which extends the scope of the directive. In the amended directive, all large companies have to perform ESEA. The proposal also introduces an audit requirement for the non-financial reports of these large companies [57]. The main purpose for this audit requirement is ensuring that the reported information is complete and accurate [37]. In other words, the final goal of the audit is to be able to provide stakeholders the *assurance* that the reports are complete and accurate. This assurance is the most important driver for companies to have their reports audited, regardless of the industry [110]. Assurance, as well as it's drivers, are discussed in the next section.

Audit processes might have other accompanying benefits. For example, many dialogues and interviews with stakeholders and interested parties are held during the audit process, which encourages better communication between parties. Incidentally, the consistency of the available information will be improved by data collection processes, which ultimately ensures better fundamental management decisions [121].

Depending on which audit is performed, the audit process might also evaluate internal systems and monitoring issues, raise questions about the materiality of the findings, and draw attention to the potential for improvement in sustainability management systems. Finally, all levels of the organization are involved in the audit process. This leads to an increased interest in sustainability management and raises awareness across all departments [122].

3.3. Assurance

The third and final phase of the ESEA process is assurance. According to several calls for action from the Federation of European Accountants (FEE), the objective of assurance of non financial reports is to provide comfort to stakeholders about the accuracy of reported data [59]. Hence, the assurance provider *assures* that the reported data is accurate. The assurance provider is able to provide assurance on the basis of the audit that has been conducted in the previous phase. The auditor's report that resulted from the auditing phase is used as input [22].

There are several definitions for the concept of assurance, some of them more narrow than others. For example, the Council of the European Union defines assurance as: 'the process of ensuring appropriate levels of confidentiality, integrity, availability, non-repudiation and authenticity' [58]. We will define assurance as: 'the process of assuring that the reported data is authentic and accurate'.

3.3.1. Levels

It stands to reason that assurance can have multiple levels. After all, an assurance provider can only be sure to a certain extend that the reported data is authentic and accurate. This relates to the depth of the assessment and affects the reliability of the assurance statement. While assurance levels originate from the financial sector, according to the International Auditing and Assurance Standards Board, the same reasoning and levels apply to ESEA engagements [81].

In theory, an unlimited number of levels of assurance could be provided that are in a spectrum from incredibly low to incredibly high. The International Standards on Assurance Engagement, as drafted by the International Auditing and Assurance Standards Board, acknowledges this: 'In theory, it is possible to provide an infinite range of assurance, from a very low level to an absolute level of assurance. In practice, it is usually not feasible to provide such fine gradations. Therefore, professional accountants usually make commitments to provide only one of two different levels: high level and a moderate level' [83]. This research follows this reasoning, distinguishing between two types of assurance: limited assurance and reasonable assurance.

Limited Assurance (also called Moderate Assurance) focuses on the plausibility that the subject matter is accurate. The assurance provider identifies 'areas where a material misstatement of the subject matter information is likely to arise' [81]. In other words, the assurance provider investigates whether the statements are worthy of believe. For limited assurance, negative assurance statements are given such as: "We have not identified material errors" or "Nothing has come to our attention" [111].

Reasonable Assurance (also called High Assurance) looks beyond plausibility and requires solid evidence to ensure the reliability of the report. The assurance provider 'identifies and assesses the risks of material misstatement in the subject matter information' [81]. In other words, the assurance provider investigates the subject matter and concludes whether it is accurate. For reasonable assurance, positive assurance statements are given such as 'the statements are free from material misstatements' [111].

It stands to reason that the workload to provide reasonable assurance is much higher than for limited assurance [76]. Moreover, according to Hasan et al, reasonable assurance is often not provided due to nature of the subject matter, unavailability of appropriate standards of evidence or performance, cost/benefit trade-offs and unavailability of appropriate supporting documentation [76].

In determining the level of assurance, certain factors are considered to be influential [76]. The subject itself is a determining factor, as some subjects are more suitable for reliable measurement than others. Another important factor are the criteria for providing assurance. Depending on the nature of the subject, some criteria provide a resource for a more reliable measurement of the subject. For example, a lower level of certainty will often be provided when the criteria are qualitative rather than quantitative [104]. This phenomena will be explored more closely in the next chapter by applying techniques such as Digit Analysis.

In addition, audit procedures are a factor that influence the level of assurance. The higher the level of assurance to be provided, the more extensive the audit procedures that must be performed [76]. Finally, the quantity and quality of the evidence collected is an important factor. If the assurance provider has obtained sufficient appropriate evidence, a higher level of assurance will be able to be offered [104]. Of course, the final level of assurance will be based on a combination of these factors.

3.3.2. Scope

Similarly to the audit procedure, assurance has a certain scope. Again, we can distinguish between a method and a data assurance. As the scope of the assurance depends entirely on the scope of the audit, this distinction will not be further discussed here. However, to further illustrate the different scopes, we will discuss a study by Mock et al. that concerned a case study into different audit and assurance processes [111].

The study demonstrated that there were different types of assurance provided, with different focuses. Mock studied 130 assurance processes used in 130 organizations worldwide between 2002 and 2004. The study shows that within these 130 organizations, different methodologies for ESEA were used. 24% of reviewers followed the AA1000AS, 18% followed international standards such as ISO, 15 % followed local standards, and 42 % did not indicate which framework was used. Organizations had different scopes. 67 % of assurance providers provided full assurance with respect to the GRI reporting categories, 16 % assured only environmental issues, while 16 % assured both environmental and social information [111].

Mock conducted an additional study in 2014, investigating the different assurance statements that were given. In 74 % of cases, the assurance provider gave a positive assurance statement. 17 % gave a negative assurance statement and 9 % gave a hybrid statement (positive and negative aspects combined) [105]. Previous academic research confirms the premise that positive assurance is perceived as the highest level of assurance, while negative assurance has the lowest level [104, 76].

3.3.3. Importance & Drivers

As stated before, the objective of assurance of non financial reports is to provide comfort to stakeholders about the accuracy of reported data [59]. For stakeholders, independent assurance means that the risk of incompleteness of data is significantly mitigated [79]. Hence, the main motivation for companies to engage in the assurance process is to improve user's perceptions of reliability of non-financial reports.

Improving perceptions of reliability of non-financial reports is, to many companies, a much needed endeavour. This might relate to the emergence of greenwashing, where companies mislead customers about their environmental performance or the environmental benefits of a product or service [46, 87]. Moreover, research suggests that the target audience of non-financial reports find that the reports often lack consistency and completeness [9].

Specifically by providing assurance statements of third parties, companies try to improve the perception of reliability of non-financial reports [45]. In 2009, Hodge et al. demonstrated that this perception of reliability is indeed improved by providing assurance statements. Moreover, the research found that the greatest increase in perception of reliability resulted from assurance statements by top tier accountancy firms such as Deloitte or Ernst & Young. It was also demonstrated that users put more trust in reasonable than limited assurance statements [79].

Simnett et al. (2009) suggest that assurance on non-financial reports does not only increase trust in information quality, but also results in 'increased stakeholder trust in the level of organizational commitment to sustainability agendas' [139]. This might be the result of users' knowledge of the costs that come with audit and assurance processes. After all, more money spent on the assurance of non-financial reporting might mean more commitment to sustainability agendas [139].

Even though assurance processes might have many benefits, not all companies voluntarily request assurance on their non-financial reports [125]. Should a company not choose to do so voluntarily, regulators might enforce assurance through legislation.

As mentioned before, in the European Union, the proposal for amendment COM/2021/189 introduced an audit requirement for the non-financial reports of all large companies

[36, 57]. The main purpose for this audit requirement is ensuring that the reported information is complete and accurate [37]. In other words, the final goal is to be able to provide stakeholders the *assurance* that the reports are complete and accurate.

The Commission sports a similar argument in the official motivation for the proposal as well: 'the principal novelties of this proposal are: [...] to require assurance of sustainability information;' [57]. The Commission also states that: '[assurance] should go a long way towards addressing the concerns of investors and other stakeholders about the reliability of the sustainability information that companies report today' [37].

For the type of assurance, the Commission adopts a progressive approach. The proposal includes an assurance requirement for limited assurance, with the option to move to reasonable assurance later. The Commission argues that, as there are no ESEA assurance standards for reasonable assurance in place yet, there is a risk of different interpretations of the concept of reasonable assurance. Should the Commission adopt ESEA assurance standards for reasonable assurance on a later date, the limited assurance requirement would automatically became a reasonable assurance requirement [37].

3.4. Computer-Assisted Audit Tools & Techniques

Computer-assisted auditing tools and techniques (CAATTs) support almost all auditing processes involving data collection and analytics. In scientific literature, multiple terms can be identified that refer to this concept, most notably computer assisted audit *tools* [13] and computer assisted audit *techniques* [18, 23]. A combination of these terms has also been proposed: computer assisted audit tools & techniques [137, 103, 10, 70]. This last term will be adopted in this research as well, as we want to investigate both the tools that are used to support the audit process, and the underlying techniques that these tools employ.

We define CAATTs as "any use of technology to assist in the completion of an audit' [24]. This (broad) definition includes, for example, spreadsheet applications (Microsoft Excel) and SQL databases, but also information analysis, artificial intelligence and generalized audit software [103, 70, 12]. In general, the use of CAATTs is positively associated with the quality of the audit report even after controlling for the big four international audit firms and internal control systems [24].

The type of CAATT that is most used by auditors is Generalized Audit software (GAS) [10, 19, 103]. GAS is a type of CAAT that is used to analyze and audit live or extracted data [44]. According to Ahmi et al (2013), GAS can be 'data extraction and data analysis software, which is designed to perform specific audit routines and statistical analysis' [10]. As the aim of this research is to design GAS functionality for the openESEA tool. Hence, in this paper, we will focus on GAS and it's characteristics.

Auditors have a number of motivations for using GAS. Research indicates that the amount of misinformation in reports is still high [84]. Errors that can lead to fraud need to be detected. A study conducted by Debreceny et al. examined the use of CAATTs in the banking industry in Singapore. The study found that GAS is frequently used in bank audits for sample extraction, identifying reactivated accounts, and verifying data completeness and accuracy [44].

Examples of GAS applications that are used in practice are Audit Command Language , CA's Easytrieve, Statistical Analysis System, Statistical Package for Social Sciences (SPSS) and IDEA. In addition, the Big Four accountancy firms also use proprietary applications such as Sustainability Assurance Methodology (SAM), Galvanize, Enablon, Helix and Canvas. In chapter X, these tools will be further elaborated upon. We will discuss what these applications entail, what techniques may be relevant, and how they are applied in an ESEA context.

3.4.1. Adoption

Research on CAATT adoption has been conducted many times in different fields and countries over the years. Within all the studies, various angles have been approached and various methods have been used.

In 2002, Banker et al. conducted interviews with auditors in an international public accounting firm and obtained quantitative panel data that explored the impact of information technology on audit firm's service production. This research resulted in some key findings, namely that the adoption of CAATTs improves individual auditor's efficiency, reduces time in preparing working papers, facilitates decision-making process, improves decision quality and business process productivity and increases firm's revenues [16].

In 2003, Braun and Davis came up with an online survey that was sent out and reached 90 legislative auditors from various states in the US. The purpose of this survey was to examine auditors' experience in two prominent GAS applications. The results revealed that auditors have a low perception of their skills in using these tools, although they perceived them as being beneficial to their work [24].

Subsequently, In 2008 and 2009, a series of studies applied well-known models to investigate the adoption of CAATTs, namely Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology [88, 92, 102, 39, 80].

As Janvrin describes in his study, the uptake and acceptance of CAATTs can be influenced by several factors. For example, one influential factor is said to be the degree to which an individual believes that using the tool will help him or her achieve superior results [88]. Moreover, auditors might perceive that the use of CAATTs will help them meet their time-budget for the audit, as CAATTs decrease the number of hours spent performing audit. In addition, ease of use has a large influence on the adoption of CAATTs [152].

The individual experience of auditors has no effect on the willingness to adopt CAATTs. Experienced auditors were expected to invest more effort in adopting CAATTs. Technical complexity and attitude of internal auditors were found to be barriers to the use of CAATTs. In short, the more complex the CAATT is, the less willing auditors are to accept and want to use it [39].

Lastly, facilitating constraints are a factor of influence. This is referred to as the extent to which an individual believes that an organizational and technical infrastructure exists to support the use of the tool [152]. In the context of CAATTs for ESEA, infrastructure may include audit firms providing appropriate CAATT resources and computer support to their employees, such as guidelines, support and/or specialized documentation [146].

The main overarching, and aligned conclusions that the previous studies have is that performance expectations, ease of use and expected decrease in number of hours spent on performing audit are significant factors in influencing CAATT adoption by individual auditors in organizations.

More recent studies have been able to reach a larger number of auditors for the questionnaires, and are generally oriented and focused on auditors in the Big Four Firms. Their overall conclusion is that budget has an important role in the adaptation of CAATTs [100, 39, 40, 18].

When budget pressure is high, the intention to adopt CAATTs is driven by the amount of expected effort involved in compliance with the expected performance [39]. Another key finding is when budget pressure is low, intention to adopt CAATT's is driven by social influence via expected higher performance [18].

Companies can influence the use of new technology by using longer-term budgets and review periods of these techniques and tools [40]. Auditors are more likely to implement new technology when the management and directors and/or partners encourages technology implementation [100]. Lastly, in an individual setting, individual auditors willing to take risks and are more likely to use technology in auditing, regardless of budget pressures [39].

4. State of the Art on Audit Method Fragments of Existing ESEA Methods

In the world in which ESEA operates, there are 65 methods gathered so far [56]. All of these methods have a different structure and approach to accounting, reporting, auditing and assuring non-financial data. This chapter will take a deeper look at state of the art on audit method fragments of existing ESEA methods and answer RQ1. In this study, 10 different ESEA methods have been investigated of which 6 are created. Some of these methods are already modelled by other students or researchers and the additional methods are chosen because there was enough documentation to create process-deliverable diagrams (PDDs) [25]. The full representation of these PDDs, contain a process (left) and a data (right) part. This way of modelling allow us to to have an overview of the related deliverables. All these included methods have been validated and approved by external experts. Afterwards two criteria were analysed and confirmed by the experts.

From the outset, it should be noted that all methods have been reviewed and modeled using their accompanying documentation. Also, where possible, experts have been consulted to validate the models. Accordingly, the terminology used to describe each method is in line with that method's documentation. While this guarantees that the models are usable and understandable outside the scope of this research, terminology differs from method to method. Therefore, similarities and differences between methods will be assessed during method comparison, which will be discussed after all individual methods.

4.1. Method Selection

In order to provide an overview of the 10 investigated methods, the following table was created. The table represents whether the method contains an audit component and whether the method is used in the professional context. The methods were chosen by examining whether there is an audit component in the ESEA method. These selected methods were validated through expert assessment and the resulting characteristics have been confirmed by them. The selection of the ESEA method is based on criteria.

The first criterion, existence of an audit component in the method, was assessed by analyzing method documentation. This activity was conducted by the researchers themselves. This assessment of the method documentation was of a preliminary nature, and as such the extensiveness of the audit component was not guaranteed.

The second criterion, expert recommendation, was assessed in collaboration with multiple experts. These experts included consultants and managers of two Big 4 auditing firms, as well as two researchers in the ESEA field. Expert recommendation was valuated on the basis of the experts having any prior experience with the method in their practice and whether this method is still being applied. We have purposely elected to make the valuation of this criterion entirely subjective, as this ensured that we could include those methods that experts in the field found to be most suitable. Finally, a brief rationale for using or not using the ESEA method is given. The resulting ten methods can be found in Table 1.

The included methods can be found in Table 1 below. The full visualization of all these methods can be found in Appendix B respectively.

Method Name	**Audit**	**Recommendation**	**Inclusion rationale**
EMAS	Yes	Not recommended	EMAS is an extension of ISO, but has never proven itself in practice.
ISAE3000	Yes	Yes, by Big Four practitioner	This ESEA method is widely used bby Big Four firms and also recommended due to the fact that this method maintains good guidelines at policy level.
GHG	Yes	Not recommended	/
NCP	Yes	Not recommended	/
B Impact Assessment	Yes	Yes, by researcher	Major actor in the domain of ESEA methods deployment.
ISO14011	Yes	Yes, by EMAS and Big Four practitioner	Used by global businesses, private enterprises, governments, and other public organisations.
AA1000	Yes	Yes, by researcher	Used by global businesses, private enterprises, governments, and other public organisations.
XES Social Balance	Yes	Yes, by researcher	It is becoming more widely known and used. The method is also constantly being expanded.
FLA-SCI	Yes	Not recommended	/
GRI	Yes	Yes, by Big Four practitioner	Used by global businesses, private enterprises, governments, and other public organisations.

Table 1: Selected ESEA methods

4.2. Method Modelling

To preserve the clarity of order and the overall picture of this study, only the first PDD out of the 10 methods will be added and described in more detail in this chapter. Other PDDs can be found in the Appendix A. The method included in this chapter is EMAS. The following figure visualizes only the fragment that applies to auditing and/or assuring non financial data.

Particularly seen in Part 1, one of the most important steps of this audit is to retrieve documentation and then to review and analyse it. Documentation is the document that is requested as evidence of the information entered. Further the conformity between the requirements and performed steps are assessed and controlled. Another important part of EMAS verification and validation is the activity 'Check reliability, credibility and accuracy of data included'. This process step compares documents and verifies the authenticity and accuracy of information entered. The activity is performed manually until today and creates space for improvement.

In Part 2 of the EMAS audit, visualized in Figure 3, we see that corrective action is required and this requires supporting documentation and evaluation of effectiveness. These particular activities of the audit are also found in other methods.

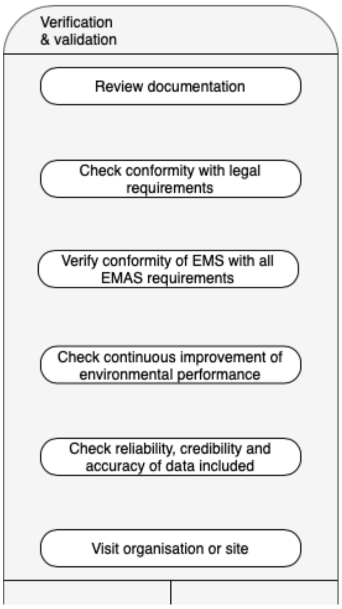

Figure 2: EMAS Audit Part 1

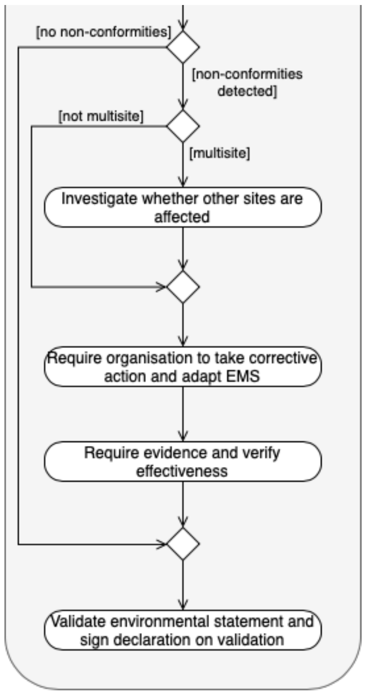

Figure 3: EMAS Audit Part 2

4.3. Development of super-method

After elaborating and creating the 6 PDDs and analyzing the, in total ten, different methods, the next step was to develop a super-method. As described in Chapter 2.7.2, the super-method is a method in which all activities from different methods are included. This means that the activities must occur at least once in a method. During the development specific activities were examined, and if they were redundant, they were generalized and translated into one activity. By applying this approach is, it became possible to visualize a super-method in a PDD, shown in Figure 4 and Figure 5. What can be seen in these three figures, is that the super-method includes 4 main Stages, including Initiate audit, Prepare audit, Conduct audit and Report audit. Additionally, it can be seen in the activity Conduct Audit, Figure 5, that four activities with the first part in the name Elicit Information, apply different tools to gather information. The process is modeled in such a way that one can choose from these information gathering techniques. It is not necessarily the case that all of them are applied, but this does not exclude them either. The super-method distinguished between documentation and evidence. Documentation is used in this study as evidence that proves that the information during the assurance of non-financial information is correct. In other words, the demonstrated information should be compared with information that supports it. Evidence, on the other hand, is information that proves that, if improvements or adjustments need to be made at an organization, these adjustments have actually been made.

4.4. Comparison of Methods

After the creation of Super-Method, an activity comparison is conducted. Within this step we create an overview how often a certain activity is performed. After finishing this comparison, the results are translated in a Generic Model. After the creation of Super-Method, an activity comparison is conducted. Within this step we create an overview how often a certain activity is performed. After finishing this comparison, the results are translated in a Generic Model.

In order to create a generic model, a threshold needs to be set how often a certain activity accurs, can be seen in column total of Table 2 and Table 3. In this Generic model, only activities with a 5 times frequency of occurrence were included in the model. The reason that 5 was taken as the threshold is to keep the most important and frequent activities. There was a consideration to take 4 as the threshold, however, the activities after which 4 was the threshold, turned out to occur only in the same methods. This did not reflect the commonality of other methods. Table 2 and Table 3 show the activities that had been used in the super-method. In the table you can see the ID of the activities, the name of the sub activities and the total number of threshold. The meaning of the different signs is explained in Chapter 2.7.4.

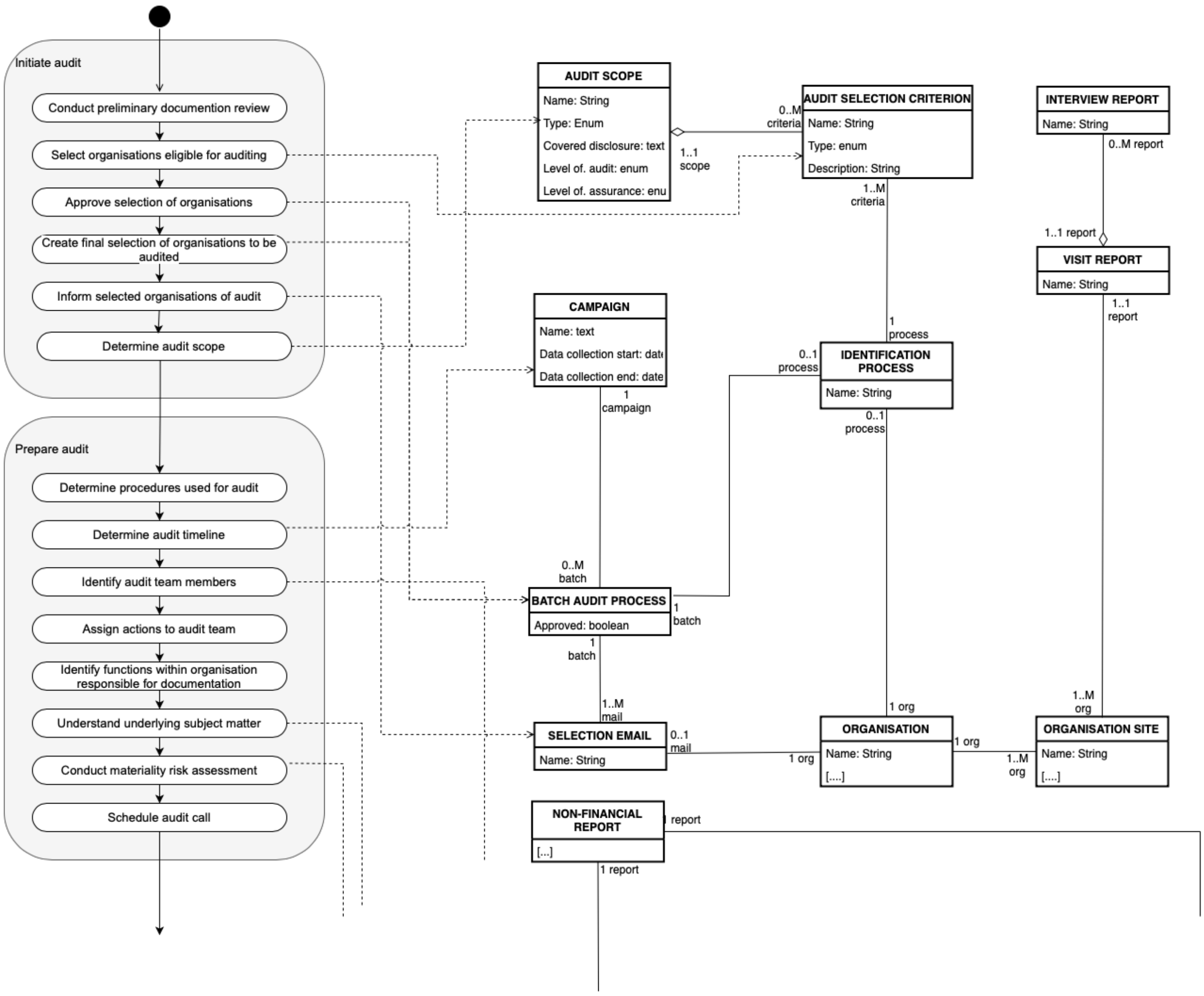

Figure 4: Super-method Part 1

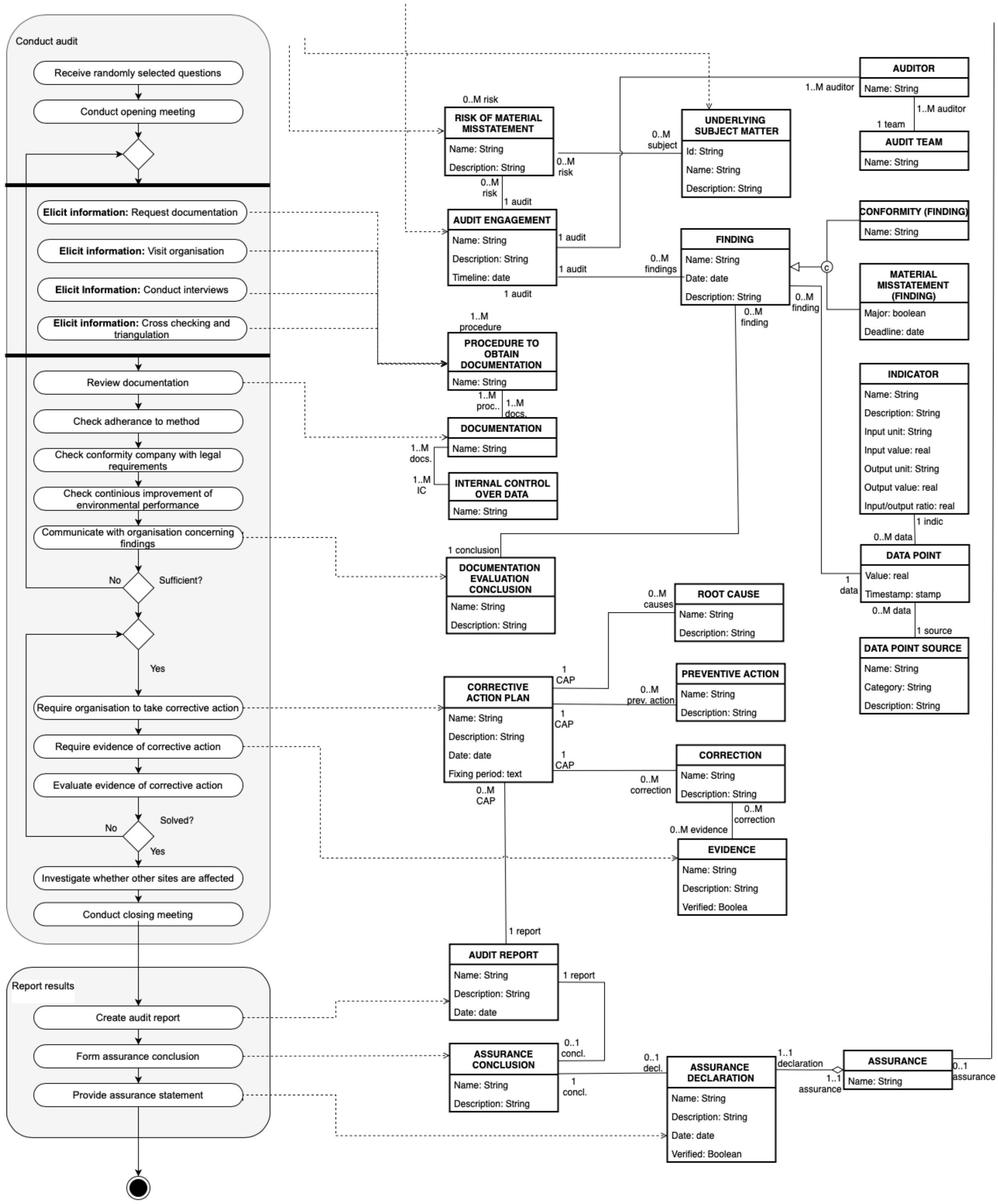

Figure 5: Super-method Part 2

ID	Activity	Sub-activity	Total	EMAS	ISAE3000	GHG	NCP	BIA	ISO14011	AA1000	XES	FLA-SCI	GRI
A	Initiate audit												
	A.1	Determine audit scope	6	=	=	=			=	=		<	
	A.2	Conduct preliminary documention review	2	=					=				
	A.3	Select organisations eligible for auditing	1								=		
	A.4	Approve selection of organisations	1								=		
	A.5	Create final selection of organisations to be audited	1								=		
	A.6	Inform selected organisations of audit	1								=		
B	Prepare audit												
	B.1	Determine procedures used for audit	4	=	>				=			<	
	B.2	Determine audit timeline	5	=	<				=	=		<	
	B.3	Identify audit team members	4	=	<				=			<	
	B.4	Assign actions to audit team	4	=	<				=			<	
	B.5	Identify functions within organisation responsible for documentation	1						=				
	B.6	Understand underlying subject matter	4	<	=				<	=			
	B.7	Conduct materiality risk assessment	5	=	=	=			=			><	
	B.8	Schedule audit call	1					=					
C	Conduct audit												
	C.1	Receive randomly selected questions	1					=					
	C.2	Conduct opening meeting	2						=			=	
	C.3.a	**Elicit information:** Request documentation	8	=	=	=		=	=	=	=	=	
	C.4	Review documentation	8	=	=	=		=	=	=	=	=	
	C.5	Check adherance to method	3	=					<	=			
	C.6	Check conformity company with legal requirements	3	=					<			>	
	C.7	Check continious improvement of environmental performance	1	=									
	C.3.b	**Elicit information:** Visit organisation	4	=	=				=			=	
	C.3.c	**Elicit Information:** Conduct interviews	2					>				=	
	C.8	Communicate with organisation concerning findings	5	>	<			>	>		=		
	C.9	Require organisation to take corrective action	8	>	=	=		<	<	=	=	<	
	C.10	Require evidence of corrective action	8	=	=	<		<	<	=	<	<	

Table 2: Method Comparison Table Part 1

ID	Activity	Sub-activity	Total	EMAS	ISAE3000	GHG	NCP	BIA	ISO14011	AA1000	XES	FLA-SCI	GRI
C	Conduct audit												
	C.1	Receive randomly selected questions	1					=					
	C.2	Conduct opening meeting	2						=			=	
	C.3.a	**Elicit information:** Request documentation	8	=	=	=		=	=	=	=	=	
	C.4	Review documentation	8	=	=	=		=	=	=	=	=	
	C.5	Check adherance to method	3	=					<	=			
	C.6	Check conformity company with legal requirements	3	=					<			>	
	C.7	Check continious improvement of environmental performance	1	=									
	C.3.b	**Elicit information:** Visit organisation	4	=	=				=			=	
	C.3.c	**Elicit Information:** Conduct interviews	2					>				=	
	C.8	Communicate with organisation concerning findings	5	>	<			>	>		=		
	C.9	Require organisation to take corrective action	8	>	=	=		<	<	=	=	<	
	C.10	Require evidence of corrective action	8	=	=	<		<	<	=	<	<	
	C.11	Evaluate evidence of corrective action	8	=	=	<		<	<	=	<	<	
	C.12	Investigate whether other sites are affected	1	=									
	C.3.d	**Elicit information:** Cross checking and triangulation	1									=	
	C.13	Conduct closing meeting	2						=			=	
D.	Report audit												
	D.1	Create audit report	9	=	=	=	=		=	<	=	>	=
	D.2	Form assurance conclusion	8	=	=	=	=	=	<	=			=
	D.3	Provide assurance statement	8	=	=	=	=	=	<	=			=

Table 3: Method Comparison Table Part 2

4.5. Generic Model

By performing all the previous steps, it was possible to develop a generic model. The development process was quite straightforward and self-explanatory. Through the activity comparison and by setting the threshold at 5, it allowed us to build a generic model. This generic model can be found in Figure 6. According to this generic model, we can conclude that the key components of this overall model are mainly two activities.

First of all conducting audit, by applying different techniques to elicit information. In this particular model requesting documentation and reviewing it is often performed in several methods. But also requiring evidence of corrective action and evaluating the evidence are important.

Second most important activity is reporting the whole audit. This includes the creation of an audit report and provision of an assurance statement. Obviously, the type of assurance and thereby the statement depend on what type of audit was conducted and what the scope discussed was. As described in Chapter 3.2.2 there is a distinction in audits and stated in Chapter 3.2.3 different scope can be agreed upon. In Chapter 3.3.1. more depth is given to the two forms of assurance.

The generic model represents the 36.36% of the super-method. These are the most important fragments from the entire audit of different methods. To support at least one method fully within openESEA, we decided to add four additional sub activities to the Generic Model. By adding these four activities in the activity 'Initiate Audit', it allowed us to support the method XES.

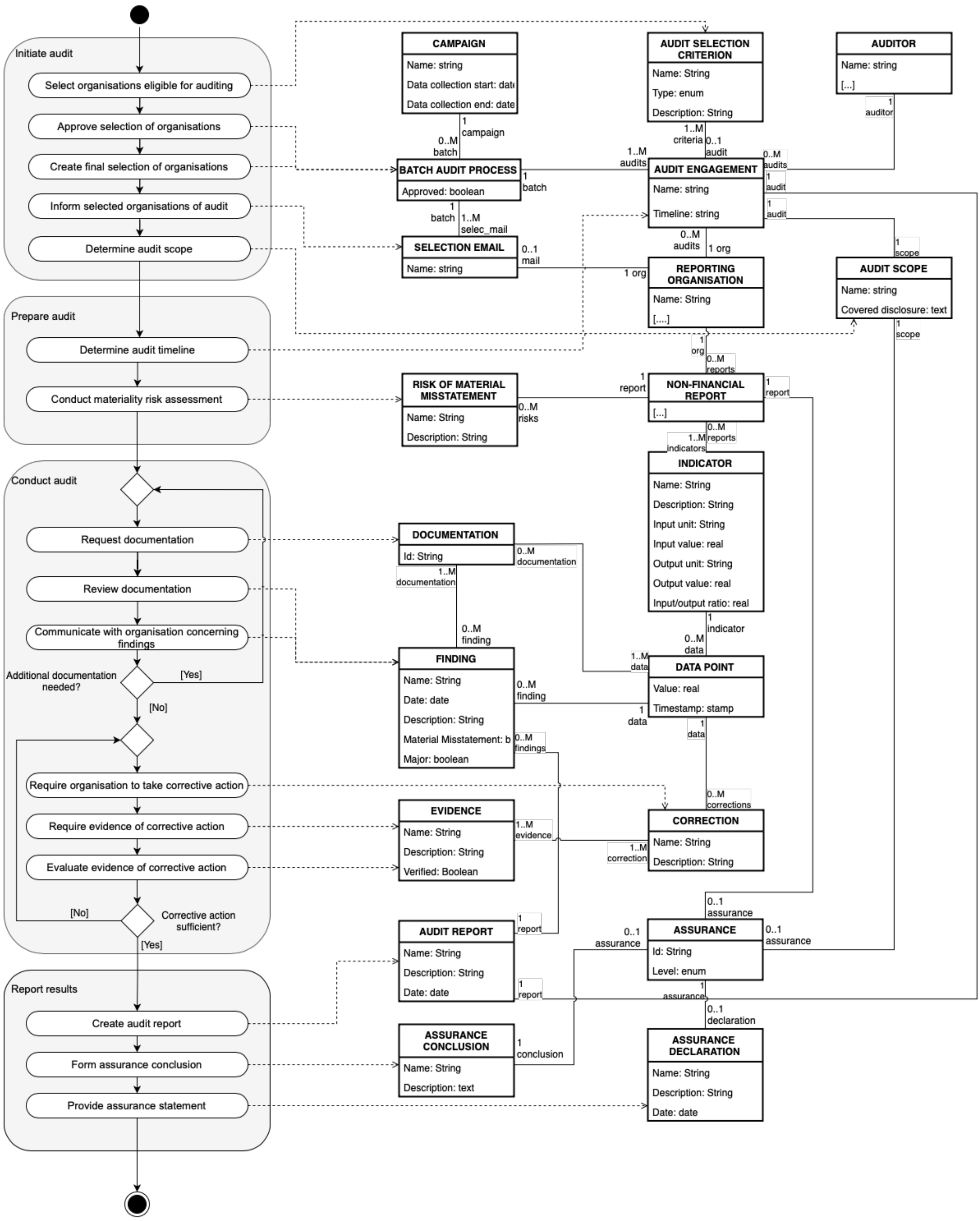

Figure 6: Generic model PDD

5. State of the art on Computer-Assisted Audit Tools and Techniques

Nowadays, almost all audit processes concerning analysis and data extraction are supported by CAATTs [5]. As stated in Chapter 3.4. the aim of these tools was first of all to support financial auditing processes. However, their scope is currently beyond this, therefore, we present CAATTs in the ESEA context.

As discussed in the background of this research, these tools commonly serve two purposes: analysing data and managing audit routines. Based on this distinction, we will identify two types of features: analytical features and managerial features. Analytical features are defined as 'features that support analysing data'. Managerial features are defined as 'features that support managing the audit process'.

For this research, we will focus on Analytical features. This chapter explores what analytical features or techniques might be useful for improving the audit method fragments and thereby answers RQ2: 'What is the state of the art on analytical features in computer-assisted auditing tools and techniques?'.

Table 4 shows the different analytical techniques chosen, based on several previous studies that have applied analytical techniques in audit context and reviewed them [7, 90].Table 4 shows the reference of other research as well as whether the technique is selected for openESEA. The selection of the techniques in Table 4 is based on whether the technique is applicable to support (accounting or) auditing as a CAATT or not. The techniques were chosen to explore more in depth based on the applicability of openESEA and associated possibilities such as data availability, current state of the tool, and the scope of this research. The subsections will elaborate on the technique and explain the chosen technique in more detail.

Technique	Reference	Selected	Rationale
Benford's Law	[17, 89, 118, 116]	Yes	Technology is usable, but not yet possible to apply. An array of numbers in the distribution being analysed need a spread across at least an order of magnitude. E.g. per organisation, one indicator needed over several years.
Anomaly detection	[64, 63, 86, 129]	Yes	Can be applied to natural numbers that have (no) restrictions. Allows the application of reasonable ranges.
Neural Network	[60, 123]	No	The presence of the data in openESEA does not allow for neural networks.
Regression	[141, 155]	No	The presence of the data in openESEA does not allow for regression.
Tree-based Algorithms	[159, 156]	No	The presence of the data in openESEA does not allow for tree-based algorithms.
Visualization methods	[7, 52, 153]	Yes	An effective means of conveying information and can be combined with other methods..
Clustering	[156]	No	The presence of the data in openESEA does not allow for clustering.
Scoring scheme	[14]	Yes	ESEA Methods use scoring indicators. By developing a scoring scheme, it is possible to see which indicators lead to certification.

Table 4: Applicable CAATTs for openESEA

5.1. Benford's Law

Many people who work with data will have heard of the Gaussian distribution. It appears regularly in all kinds of natural phenomena or observations. However, there is another natural law regarding numbers that is less well known but is all around us and is increasingly used to fight fraud or make predictions. In 1997, this method was more widely researched and applied for use in audit processes to detect fraud using Benford's Law as an aid in analytical solutions. [118].

Benford's Law, also known as Phenomenon of Significant Digits or First Digits, describes the frequency distribution of the first digits (or numerals) of numbers found in large data collections or series of records from various sources [17]. This frequency distribution does not follow a uniform or normal distribution, but exhibits a ranked order where the digit '1' occurs most frequently, followed by '2', '3', '4' ... '9' which occur less frequently respectively.

It is based on the premise that the distribution of digits in sets of natural numbers is not random, but follows a predictable pattern. It refers only to 'natural numbers.' There is a distinction that should be highlighted in terms of the definition. How the distinction is made in this study is as follows:

- Natural numbers are numbers that are not ordered or placed in a particular numbering order and are not human manipulated or being created from a random number system.
- Non-natural numbers are systematically designed or manipulated to convey information that mitigates the natural nature of the number (e.g., phone numbers, zip codes).

Although this phenomenon received its name from Benford [17], Benford's Law is based on an earlier study by Simon Newcomb that was conducted in 1881. Simon Newcomb noticed this from logarithmic tables where the first pages of the books were often more used and worn out. From this observation, Newcomb devised a mathematical formula that calculates the probability that number d will occur as the first digit in the set of numbers D.

Benford's empirical test had an approach to test the first digit frequencies in 10 lists of small numnbers (approximately 3.000 observations) and 20 lists of large numbers (approximately 20.000 observations). These lists covered data that both were weakly dependent and independent. Weakly dependent lists covered tables of mathematics from several engineering handbooks and tabulations of weights including different physical constants. Independent lists were taken and constructed from different sources such as drainage areas of random rivers and numbers numbers appearing in a specific issue of Reader's Digest. After this discovery, he hypothesized that data which occurs in a natural order, not being manipulated, should form a geometric sequence. He then formulated through the use of internal calculus the expected digital frequency for the first and second digits and it's combination.

The following rule is defined by the following logarithmic function, which describes occurrence of first digits in a large number of data sets:

$$Prob(D_1 = d_1) = log\Big(1 + \frac{1}{d_1}\Big) : \qquad d_1 \in [1, 2, ...9] \tag{1}$$

Equation 2 calculates the occurrence of the second digit in a set of numbers of D. Notice that the digit 0 does occur in this set of numbers.

$$Prob(D_2 = d_2) = \sum_{d_1=1}^{9} log\Big(1 + \frac{1}{d_1 d_2}\Big) : \qquad d_2 \in [0, 1, 2, ...9] \tag{2}$$

Finally, we have the combination of both the equations, in Equation 3, where the frequency of the first two digits is calculated.

$$Prob(D_1 D_2 = d_1 d_2) = log\Big(1 + \frac{1}{d_1 d_2}\Big) : \qquad d_2 d_2 \in [10, 11, ...99] \tag{3}$$

Benford's empirical research shows that the odds are not random. On the contrary, there is a 17.6% chance that the first digit will be a '2' and that there is a 30.1% chance that the first digit will be a '1'. In Table 5 we can see the percentages of occurrence of a specific digit.

Digit	**1st**	**2nd**	**3rd**	**4th**
0		.11968	.10178	.10018
1	.30103	.11389	.10138	.10014
2	.17609	.10882	.10097	.10010
3	.12494	.10433	.10057	.10006
4	.09691	.10031	.10018	.10002
5	.07918	.09668	.09979	.09998
6	.06695	.09337	.09940	.09994
7	.05799	.09035	.09902	.09990
8	.05115	.08757	.09864	.09986
9	.04576	.08500	.09827	.09982
Sum	1	1	1	1

Table 5: Benford's Law: Expected Digital Frequencies

In order to visualize how the frequency of the distribution would look like of the first digit, a scatterplot is created and shown in Figure 7. On the y-axis the percentage of the probability that digit 1 accurs in a dataset, and on the x-axis the set of numbers from 1 to 9 corresponding to the probability of occurrence is visualized. In Appendix C a histogram is created to showcase this distribution.

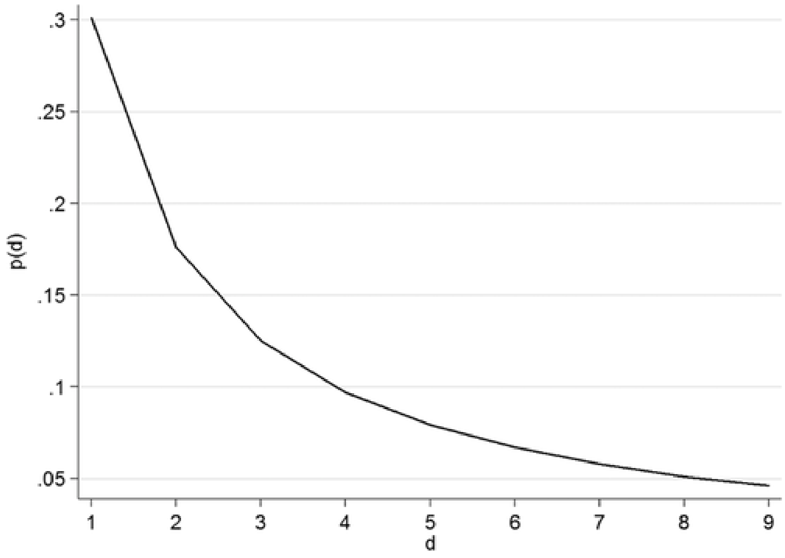

Figure 7: Frequency Distribution according Benford's Law

5.1.1. Mean Absolute Deviation

In order to determine the extent to which a data set conforms to Beford's Law, Mean Absolute Deviation (MAD) is often used. This goodness-of-fit test often used in Digital Analysis, is calculated by summing the absolute deviations and then dividing by 9 (since there are nine 1st digits) [117]. MAD consists of a dependency on the size of the considered data set and the formula of Mean Absolute Deviation can be seen in Equation 4 or Equation 5; where N is the size of the sample; x_i is the value of the sample considered and $\overline{x}$ or $m(X)$ is the value that is expected. The purpose of the absolute symbol is to ensure that the deviation that is present always gives a positive sign to the value, regardless of whether it is positive or negative.

$$MAD = \frac{1}{N}\sum_{i=1}^{N}\Big|x_i - m(X)\Big| \tag{4}$$

OR

$$MAD = \frac{\sum_{|x_i - \overline{x}|}}{n} \tag{5}$$

The formula gives a value as a result, where you can deduce that the higher the MAD, the greater the mean deviation between the actual and expected ratio. That is, if the MAD represents a large number, Benford's Law is less corresponding to the data and may lead to data manipulation or fraud [118]. In other words, the higher the MAD, the larger the average difference between the observed and theoretical distributions.

Nigrini and Drake initially developed in 2000 critical value ranges, which apply to the first, second, and combined first-two digits [117]. 12 years later, Nigrini modified the conformity ranges in a subsequent study. Nigrini made this adjustment based on the fact that small data sets can be prone to false positive errors when the results conclude nonconformity from unsubstantiated data [116]. He adjusted the conformance ranges to increase the effectiveness of the calculation. The adjusted ranges of critical values for the mean absolute deviation are shown in Table 6.

Conformity Range	**First Digits**	**Second Digits**	**First Two Digits**
Close conformity	0.000-0.006	0.000-0.008	0.0000-0.0012
Acceptable conformity	0.006-0.012	0.008-0.010	0.0012-0.0018
Marginally acceptable conformity	0.012-0.015	0.010-0.012	0.0018-0.0022
Nonconformity	Above 0.015	Above 0.012	Above 0.0022

Table 6: Mean Absolute Deviation Critical Value Ranges

In Nigrini's research, he states that the first digit test is generally used as an initial high-level test of reasonableness. The second digit test is usually used as a second high-level test of reasonableness. And the objective of the first-two digits test is to look for spikes [117].

5.1.2. Chi-Square Test

Another test that is often used to test conformity to Benford's Law is Chi-Square test. To compare a set of actual results with the expected results, Chi-square X^2 may be interesting to calculate and see the differences between MAD and X^2. In this research chi-square is considered to compare a set of survey responses with the expected results. The expected result is that the survey responses conforms to Benford's Law.

$$X^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \tag{6}$$

O_i = Observed values
E_i = Expected Values

According to Nigrini, Chi-square test suffers from the problem that as the number of records increases, the calculated chi-square will almost always become larger than the critical value. This subsequently leads to the conclusion that the data is not consistent with Benford's Law. In the same book Nigrini claims that, the chi-square test has since been relegated to the second favorite test for conformity in favor of the MAD test[116]. However, in this study both test methods will be calculated to observe actual difference. Important to consider in examining this method, is the fact that this technique focuses particularly on analyzing series of natural numbers that have no limited range such as Likert scale.

5.2. Anomaly detection

Anomalies are instances which are to some extent unusual and do not seem to fit the general patterns currently present in the dataset [64, 63, 86]. Such instances are commonly also referred to as outliers, novelties or deviant observations [64]. An outlier is simply a data point that is significantly different from other data points in the given dataset [63]. Anomaly detection is simply the process of analyzing the dataset to find these instances. Anomaly detection can be used for various purposes, such as fraud detection, data improvement analysis, information security management, process and system monitoring, and data cleansing prior to training statistical models [64, 86, 129]. From the study of [63], an overview was retrieved that gives a theoretical, in-depth and concrete outline of the general typology of anomalies into the types of anomalies that can be found in Figure 8.

		Types of Data		
		Continuous attributes	**Categorical attributes**	**Mixed attributes**
Cardinality of Relationship	**Univariate** Described by individual attributes (independence)	Type I Extreme value anomaly	Type II Rare class anomaly	Type III Simple mixed data anomaly
	Multivariate Described by multidimensionality (dependence)	Type IV Multidimensional numerical anomaly	Type V Multidimensional rare class anomaly	Type VI Multidimensional mixed data anomaly

Figure 8: The typology of anomalies, took from [63]

The typology uses two dimensions, each of which describes a fundamental underlying part in the nature of the data, to distinguish between the types of anomalies [63]. The top dimension is representing the types of data involved in characterising the behavior of the instances. It refers to the data types of attributes (i.e., indicators or variables) that play a role in the anomalous nature of a instance and thus must be treated appropriately during the analysis in order to be detected. The data types are: Continuous, Categorical and Mixed.

- **Continuous**: Continuous data considers the variables that represent the deviant behavior where the data is all numerical in nature. Examples of such variables are weight, height, temperature, and age.
- **Categorical**: Categorical data includes the variables that represent the deviant behavior, all of which are codes or class values. This contains binomial and multinomial attributes. Examples of such variables are gender, eye colour, and hair type.
- **Mixed**: This type includes variables that captures the deviant behavior which are

both continuous and categorical. In this type at least one attribute of each type is present.

The second dimension is the cardinality of the relationship and reflects the degree to which the various attributes relate to each other in describing deviant behavior [63]. Regardless of whether individually or collectively, these attributes are responsible for the deviant instance:

- **Univariate**: Besides being in the same subset of the same dataset, there is no relationship between the variables to which deviant behavior of the deviant instance can be attributed. Thus, to detect the anomaly, its attributes should be examined separately.
- **Multivariate**: The deviant behavior of the anomaly lies in the relationships between its variables. Thus, the anomaly simply cannot be found by examining the individual variables separately. The variables need to be collectively examined to take into account their relationships, i.e., the combination of its values. An example of relationships includes i.e. correlations or associations.

However, this research will not focus on all the types. Type 1 has been selected due to the availability of the present data and attributes from openESEA. Type I is an extreme value anomaly that implies an instance with an extremely low, high, or alternatively uncommon value for one or more individual numerical attributes [63]. Conventional univariate statistics typically has methods for finding this type, i.e., using a measure of the central tendency plus or minus 3 times the standard deviation (Empirical Standard) or the absolute median deviation [64, 29]. These instances are simply 'outliers' because they lie in an isolated range of the numerical space. Given that this method is very powerful for detecting values that are outside the empirical standard, and provides the freedom and opportunity for the auditor to insert their own reasonable ranges, among other features, we will explore how to implement this CAATT in openESEA.

5.3. Visualization methods

Visualization may refer to the conveniently understandable presentation of data and to the methodologies that convert complex data(sets) into clear visual patterns [7]. This allows stakeholders to perceive the complex patterns or relationships that emerge into understandable graphs. Visualization is best used to demonstrate complex patterns through a clear presentation of data or features [52]. There are many ways to visualize data, however, according to [158] a box plot, is more than a substitute for a table. It is often being used as a tool in exploratory data analysis to improve the reasoning about quantitative information. It also is referred as a five-number summary.

As shown in Figure 9, the graph utilises different aspects to present the distribution. The boxplot displays the median. The median is the central number of a set of numbers. This means that there are 50% of the numbers under the median and 50% are above the median. Then, once again, those parts are separated into two as well. Thus, in the boxplot we can see where the 1st 25% of the numbers are located, the 2nd 25%, the 3rd 25%, and the 4th 25%. They are called quartiles because the boxplot is divided into four parts. In addition to the quartiles, there are two other important numbers in the graph, which are the minimum and the maximum [153]. Information can be presented using various visualisation methods. There are forest plots, scatter plots, line charts, bar charts, combinations of all, and much more. However, not everything is applicable for visualising the information for the auditor. In this case, it is important to know how much an organisation differs from others. By using this visualisation method, the two techniques, Anomaly Detection, described in the section above, and Boxplot can function jointly given that they share common characteristics. Incidentally, one was chosen to provide openESEA with the possibility of presenting information during the audit.

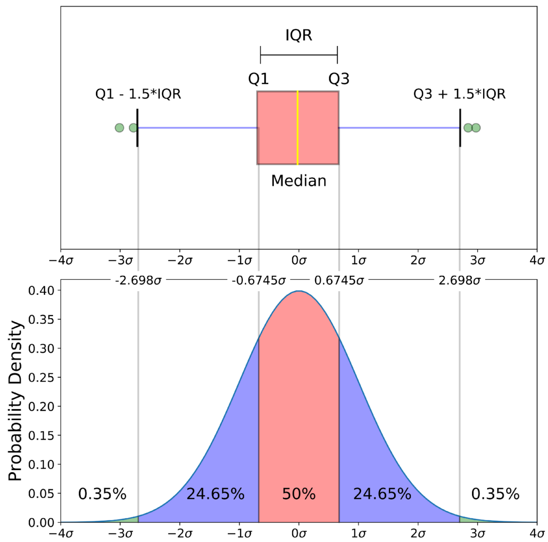

Figure 9: Different parts of a box plot, took from [66]

5.4. Scoring scheme

Scoring scheme is a technique that uses a specific ESEA method. The considered method has two types of indicators, namely the direct and indirect indicators.

The direct indicators are indicators that are not calculated by formulas but are added and stored directly. In openESEA this is done through a completed survey. Then, we have the indirect indicators. These can be calculated by applying formulas that make use of other direct and indirect indicators. They are not passed through the survey and this also explains the name 'indirect'. Each indirect indicator has its own formula with its own fixed weights to determine how heavily an indicator is taken into account in the calculation.

Indirect indicators can be divided into three types, namely performance, scoring and certification type. The performance type measures how well an organization is performing in terms of ESEA. These indirect indicators can be salary gap between men and women, percentage of dependence on grants and ratio of management positions over the total staff.

The scoring type is an indicator from which a score results. In some formulas, performance indicators are used to calculate score indicators. Finally we have the certification type. This is usually one indicator that makes a calculation, listing all the scores of indirect indicators to find out whether (or not) a company deserves a certificate.

When developing a scoring scheme, various values will have to be taken into account. As mentioned earlier, an indicator formula is provided with a weight. This weight should not be confused with the final score of an indicator, nor with the value an indicator has in the formula in relation to the overall score.

Therefore, absolute and relative value should be used. While the absolute refers to the weight that the indicator has in relation to the overall score, the relative value is based on the given weight of an indicator. If these weights are calculated correctly, we want to know which companies are just on the edge of receiving a certificate and which indicators are the ones that might cause the organisation not to get a certification. This technique will therefore recommend the indicators for the audit. Technique such as scoring scheme is also being used by organisations like B Impact Assessment [14] and could perhaps be called differently.

Figure 10 depicts how the levels within a formula are broken down and both the absolute and relative values are determined. Figure 11 technically shows the same although it also calculates what the impact of an indicator can be when the absolute score of its indicator is taken into account.

```
level 1: environmental_impact_score: 1 * 0.4
--------------
level 2: recycling_ratio_score: 0.4 * 0.5
level 2: renewable_ratio_score: 0.4 * 0.5


level 1: gender_equity_score: 1 * 0.3
--------------
level 2: gender_decision_making_ratio_score: 0.3 * 0.2
level 2: gender_pay_gap_score: 0.3 * 0.45
level 2: gender_ratio_score: 0.3 * 0.35


level 1: workplace_quality_score: 1 * 0.3
--------------
level 2: average_employee_satisfaction: 0.3 * 0.4
level 2: decision_making_ratio_score: 0.3 * 0.3
level 2: public_salaries_score: 0.3 * 0.3



[{'absolute': 0.4, 'indicator': 'environmental_impact_score'},
 {'absolute': 0.3, 'indicator': 'gender_equity_score'},
 {'absolute': 0.3, 'indicator': 'workplace_quality_score'},
 {'absolute': 0.2, 'indicator': 'recycling_ratio_score'},
 {'absolute': 0.2, 'indicator': 'renewable_ratio_score'},
 {'absolute': 0.135, 'indicator': 'gender_pay_gap_score'},
 {'absolute': 0.12, 'indicator': 'average_employee_satisfaction'},
 {'absolute': 0.105, 'indicator': 'gender_ratio_score'},
 {'absolute': 0.09, 'indicator': 'decision_making_ratio_score'},
 {'absolute': 0.09, 'indicator': 'public_salaries_score'},
 {'absolute': 0.06, 'indicator': 'gender_decision_making_ratio_score'}]
```

Figure 10: Scoring scheme output in Python

```
impact = 8.338090497737557 - 0.4 * 14.0.
environmental_impact_score has an impact of 5.6000000000000005 on the total score(8.338090497737557), corrected total score: 2.7380904977375566!

impact = 8.338090497737557 - 0.3 * 4.526968325791855.
gender_equity_score has an impact of 1.3580904977375565 on the total score(8.338090497737557), corrected total score: 6.98!

impact = 8.338090497737557 - 0.3 * 4.6.
workplace_quality_score has an impact of 1.38 on the total score(8.338090497737557), corrected total score: 6.958090497737557!

impact = 8.338090497737557 - 0.2 * 20.0.
renewable_ratio_score has an impact of 4.0 on the total score(8.338090497737557), corrected total score: 4.338090497737557!

impact = 8.338090497737557 - 0.2 * 8.
recycling_ratio_score has an impact of 1.6 on the total score(8.338090497737557), corrected total score: 6.7380904977375575!

impact = 8.338090497737557 - 0.135 * 11.25.
gender_pay_gap_score has an impact of 1.51875 on the total score(8.338090497737557), corrected total score: 6.819340497737557!

impact = 8.338090497737557 - 0.12 * 4.0.
average_employee_satisfaction has an impact of 0.48 on the total score(8.338090497737557), corrected total score: 7.858090497737557!

impact = 8.338090497737557 - 0.105 * -1.5300904977375565.
gender_ratio_score has an impact of -0.1606595022624434 on the total score(8.338090497737557), corrected total score: 8.498750000000001!

impact = 8.338090497737557 - 0.09 * 10.
public_salaries_score has an impact of 0.8999999999999999 on the total score(8.338090497737557), corrected total score: 7.438090497737557!

impact = 8.338090497737557 - 0.09 * 0.
decision_making_ratio_score has an impact of 0.0 on the total score(8.338090497737557), corrected total score: 8.338090497737557!

impact = 8.338090497737557 - 0.06 * 0.
gender_decision_making_ratio_score has an impact of 0.0 on the total score(8.338090497737557), corrected total score: 8.338090497737557!
```

Figure 11: Scoring scheme output in Python

6. Treatment Design

This section provides an explanation about the tool including the meta-model and its graph. Further, the requirements and user stories will be explained, followed by the concrete changes that have been made to openESEA. This chapter will thus answer RQ3: How can the fragment of the audit process in openESEA be improved or(semi)-automated?

6.1. OpenESEA

OpenESEA is an extendable, model-driven, open-source, web-based tool, initially developed for academic purposes by Niels Bik in 2018 during his Masters' thesis [151]. Prior research have worked on openESEA and extended the tool. For instance, several additions have been made to the tool such as [96] researched and developed a feature to specify and deliver stakeholder surveys, [47] researched and developed a feature to specify and automatically generate infographics that present the account results and [132] layed all the groundwork for the development of an extendable openESEA tool. In Figure 12 the metamoodel of openESEA is presented. The figure depicts the openESEA metamodel. The metamodel contains concepts/metaclasses and their relationships.

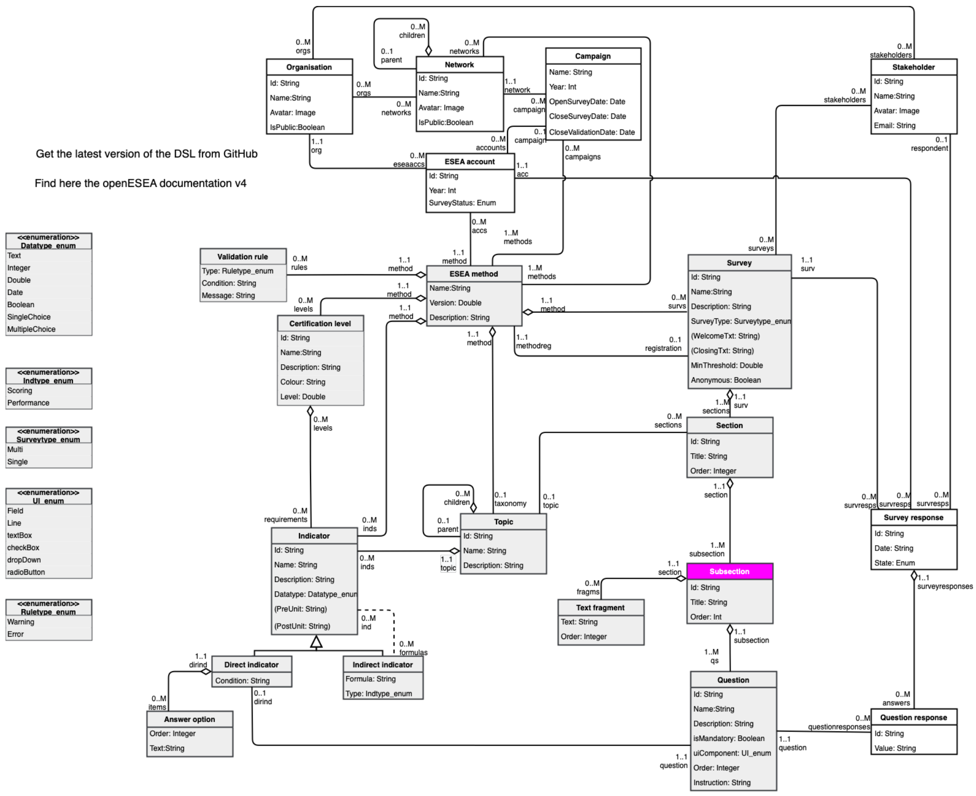

Figure 12: openESEA meta model v4.0 took from [56]

6.2. Requirements and user stories

Before the extension of the analytical features to openESEA, we need to identify what is needed for the extension. The input of these requirements is created by having multiple discussions with the research team and investigating the generic model. We were able to create requirements with corresponding user stories by studying the activities of the generic model. Some examples of the most crucial user stories for the functionality are stated below. User stories define the essential aspects of the requirements by describing who may perform the activity, what is being expected, and, optionally, why it is essential. A full definition and explanation how to create user stories can be found in a book written by Cohn [33].

As a <type of user>, I want <goal>, [so that <some reason>].

The first part is the 'type' of user who is supposed to execute the user story. In openESEA there are 7 different users. Appendix D provides full description of these user stories and their roles. All users have their own permissions and responsibilities. The second part of the user story describes the 'goal' what the user should be able to do. The third, and final, part of the user story contains the reason why a goal needs to be achieved or executed for the user. The 'reason why' a goal needs to be achieved is not essential and will not be formulated in every user story. But however, a description can be found in Tables 7, 8 and 9 that include the reasoning.

As a network administrator, I want to receive recommendations on organisations that have to be audited.

As a network administrator, I want the recommendations to be based on anomaly detection.

As a network administrator, I want the recommendations to be based on borderline certification score.

As a network administrator, I want the anomalies to be presented through a visualization method.

Complete list of requirements and user stories including 42 user stories divided in 5 epics can be found in Appendix D.

6.3. Improving the fragments of the audit process by analytical features

After examining the generic model, and looking at the process steps in openESEA, it occured that mainly the activity in the generic model 'Conduct Materiality Risk Assessment' can be improved by CAATTs. The previous user stories, combined with the openESEA meta-model have lead to an potential improvement. Based on the user stories and activities from the Generic Model, Figure 13 is created. Figure 13 shows the v4.0 metamodel of openESEA with the added audit extension.

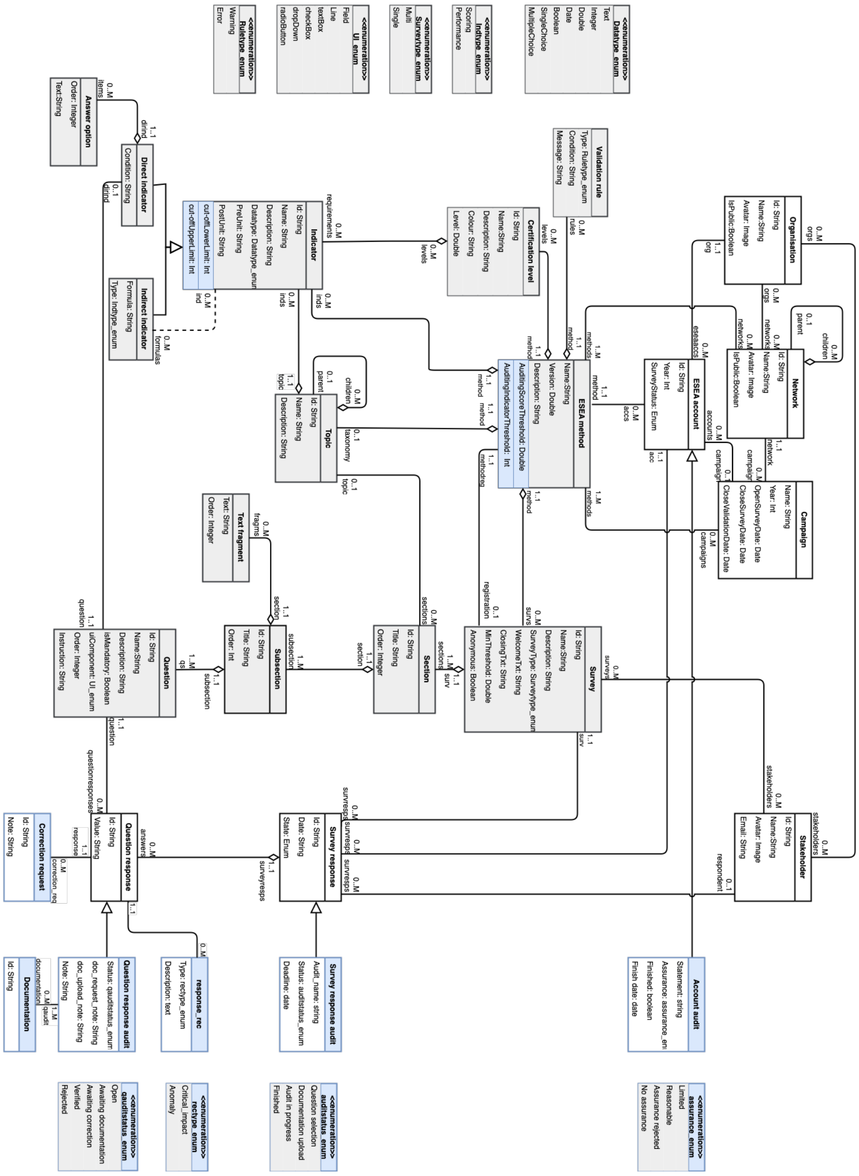

Figure 13: openESEA metamodel including audit extension

6.3.1. Concepts

V4.0 of the meta model was used since the grammer of the tool was up to date. The new meta model includes all the grey concepts from v4.0 as well as their relationships, where the light blue concepts and attributes including their relationships belong to the extension of the audit functionalities. The grey colour indicates which concepts are supported by openESEA as by the Design Specific Language. The white colour of the concepts represents that something must be initiated by the tool itself. In other words, this means that survey answers need to be completed and filled in, or the survey even needs to be created before being sent to organizations.

6.3.2. Indicator

The improvement of the audit process provides the auditor recommendations over specific indicators. This will simplify (and semi-automate) the auditor's work by no longer having to manually search and detect anomalies. By having a look at Figure 13, we can see that two attributes have been added to the concept Indicator, Cut-OffLowerLimit and Cut-OffUpperLimit. These two variables are meant to calculate the anomalies for each indicator. If the auditor decides, while creating a method, to assign a value to these variables, it will be possible to recommend an indicator for the audit if the value falls outside the ranges. This is called reasonable ranges as the auditor van decide themselves what these ranges are. If no value is assigned to the Cut-Offs, an Empirical Standard range will always be calculated and any value outside that range will be recommended. A visual example for entering the values into the method specification is shown in Appendix D, Figure 27.

6.3.3. ESEA Method

The scoring scheme provides a list of indicators. These indicators have such a weighting that have triggered organizations to just receive a certification (or not). By using the variable AuditingScoreThreshold, the auditor can assign a value to the threshold of certification. This threshold is meant to create a boundary for the organisations that just qualify or do not qualify for certification. In addition, the auditor can fill in AuditingIndicatorThreshold whereby a calculation is made as to how many indicators should just make the trigger. However, it is not yet optimally developed. If the auditor finds a organization with many recommendations, the auditor may decide to audit that organization more carefully. Furthermore, the auditor has the possibility to find out the direct indicator to audit through the scoring scheme. In this way the auditor may decide to request documentation for the specific underlying direct indicator. A visual example for entering the values into the method specification is shown in Appendix D, Figure 28 and Figure 29.

Other light blue concepts: For the information about the additional light blue concept and their attributes, please refer to the research by Ties et al [150].

7. Treatment Validation

This chapter answers RQ4.1: 'How do experts assess the audit functionality of openESEA?' and RQ4.2: 'How well does the analytical added functionality of openESEA perform?'. To answer these questions, an expert assessment was conducted with external and internal experts. The experts were invited to validate the tool and its functionalities. In addition, a mini experiment was conducted to test out the functionality and usability.

7.1. Expert assessment

A total of two expert assessments were conducted at different phases of the research during the development. These assessments were examined in addition to the weekly meetings with the research group. The two expert assessments were conducted internally within the research group with experts in the field as well as with external stakeholders. The first validation, was held with two internal experts. This assessment covered the progress of the development as well as the outcome.

During this validation, the auditing process in openESEA was partially followed and the analytical features were assessed. Through this validation some specific notes emerged and there was opportunity for improvement of the functionalities.

One of the improvements was that the scoring scheme did not display the parent of the indicator. There was a 'blackbox' that did not reveal which direct indicators should be further selected for the audit, nor which parent indicators were related to the recommended indicator.

Subsequently, it was found that the presentation of the table to select direct and indirect indicators contained too much unnecessary information (e.g. level, impact) and was in fact confusing and overwhelming.

After processing the emerged comments during the internal validation and optimizing the functionalities, an external expert assessment was conducted with experts from XES. During the validation questions such as 'What are features you are missing?' and 'Would you see yourself using this tool?' were asked and answered positively. Both Internal and External experts have showed a positive reaction to the tool. Thus, the validation reached a positive end and the experts were satisfied with both the process of the audit and the functionalities based on their usability.

7.2. Evaluation

This evaluation implies that for the introduced analytical features specific goals must be achieved as a result. For this evaluation, we want to answer questions and see if the functionalities work as they are intended. The questions we want to answer during the evaluation are: Will the anomalies be properly detected based on the reasonable ranges input from the auditor? Does the anomaly detection also calculate the extreme values (3 times standard deviation) if no values are entered for the cut-offs? Does the scoring scheme detect the direct indicator and present the critical values based on the underlying score?

First, we want to know what happens when the assigned cut-offs are not used, also see Figure D. The intended goal here would be that for any indicators that have not been assigned a cut-off value, the indicator value would not be seen as an anomaly. We tested this on an indirect indicator 'Gender pay gap' and a direct indicator 'Woman staff'. Both gave correct results. Since the number of recommendations changed we verified if they were modified in the selection of the survey. This also gave a positive outcome.

We then assessed the performance of the scoring scheme with different certification scores and thresholds. When applying a threshold of the certification score, that is close to the overall score, only those indicators that had a weak impact were detected. In this way it was possible to find out which indirect indicators led to an organization just getting certified and are critical for further audit.

By having a look at Figure 31 we can find the critical values, as well as their parent indicator and the allocated threshold. Finally, we look at the anomaly detection technique to see if the cut-offs per indicator are calculated as 3 times standard deviation (empirical standard) of its own value, if no value for the cut-off is entered. Also this calculation has been executed on each indicator properly.

8. Discussion

8.1. Limitations

We acknowledge that there are a number of limitations that may affect the validity of the study we conducted. In this subsection the limitations of this research will be discussed. During the research, and mainly during the phase of implementation, not all methods could have been applied.

For the generic model, a threshold of 5 is chosen. As earlier explained, the rationale to this decision is to support at least one method out of ten. The ESEA method that is fully supported by the generic model is XES. Clearly, openESEA could support multiple audit fragments of different ESEA methods. However, the threshold to accomplish this would have to be higher than 5. Moreover, in the best scenario, a validation by all stakeholders should be performed for the generic model to investigate how all audit fragments from all possible methods can be supported. The generic model has not been ultimately validated with all experts of various methods, but rather with one expert of XES method and with the research group. In the ideal situation, the validation could have been performed with all stakeholders.

Regarding the extension of the analytical features for audit in openESEA, we have introduced the first proposal as to where these CAATTs could be applied on. First of all, the implementation of Benford's law is limited by the availability of data. To perform Benford's Law in its completeness and correctness, it is required to use the same indicator. This could be the same indicator of same company over multiple years. By doing that, the probability of the distribution of the digits will be targeted on one case, and thus, it will allow the auditor to infer whether the company's data is correct and has no anomalies in its distribution.

In addition, this research focused on introducing analytical features to openESEA, but there is no comparison yet made on different anomaly detection methods. This comparison could be made by experimenting and combining other methods stated in Table 3.4 and ultimately adopt the best anomaly detection method.

Furthermore, little documentation could be found on the scoring scheme, making it more complicated to substantiate scientifically. Giving the fact that this technique has been implemented for the first time in openESEA, improvements will undoubtedly be possible. The choice was made to support the presentation of data with a visualization technique. The reason for displaying a boxplot per indicator was due to the fact that it fits well with the anomaly detection technique that has been developed. However, there are many other possibilities and a boxplot is not the only way to present quantitative information.

Lastly, we have chosen to look specifically at methods that can easily support ESEA audits, but methods that support auditing financial data were examined relatively less. However, numerous studies have been conducted that detect fraud on financial statements by making use of Isolation Forest, Clustering and Classification. Further studies could examine how these techniques could be applicable to ESEA as well as integrated in openESEA tool.

8.2. Future work

As there are many analytical features that can be explored and openESEA is constantly evolving, therefore, many new opportunities come to the surface to expand the audit process within openESEA in the future. The groundwork of implementing the audit process and improving it by applying analytical features, allows us to propose the direction where future work may lay.

Starting with the first implemented feature, Benfords Law. In future research, this specific CAATT could be applied on one company, where one indicator is taken into account over multiple years, i.e. ten or more years. Moreover, the indicator must be of a natural number that are not ordered or placed in a particular numbering order and are not human manipulated or being created from a random number system.

The second implemented feature is the anomaly detection with reasonable ranges. During the implementation different anomaly detection algorithms were considered, but no clearly comparison with results of these techniques is made. The anomaly detection is now performed on univariate, continuous attributes (Type 1), for future work other types could be considered to research more in depth, for example Type 4 by applying isolation forest or regression analysis.

The third feature that is implemented in openESEA is scoring scheme. The current state of the scoring scheme is focusing on whether a organisation receives the certification, which has a specific threshold. The current state of the technique is based on the (not) obtaining of certification whereby the indicators with the highest absolute value are deducted from overall organization score. However, future work could explore whether different levels of certification (i.e. bronze, silver and gold) and different thresholds are possible for certification. In addition, another calculation could be made to see which other indicators are just on the edge of receiving certification or not. The intention in the future is to generate reports using openESEA. This is a functionality that still needs to be built in and implemented. Once that is accomplished, it could be explored how Natural Language Processing might be used to compare the reports with the supporting documentation. The first step towards this could be to compare the keywords of both reports. For example, one could look at CO2 omission, number of people employed and the gender gap in the information entered in openESEA and the annual reports.

In this research, due to time constraints and scope, it was decided to evaluate how the tool performs. However, in the future an experiment could be conducted with two different groups, experimental and controlled group [15]. The experimental group could test the tool by going through the audit process with and without recommendations [77]. The independent variable could be with and without the analytical functionalities. This group is exposed to changes in the independent variable being tested. In addition, a controlled group would be separated from the rest of the experiment such that the independent variable being tested cannot influence the results. In addition, factors influencing the result may also be included. For example, during the experiment, the influence of the change in the certification threshold can be examined and what the auditor's intuition is in applying reasonable ranges.

All data that we have so far considered is data that should be provided by a single stakeholder, most often the accountant of the reporting organisation. However, in some cases, ESEA data is based on data of multiple stakeholders at once. One might consider, for example, a employee survey in which all employees of an organisation indicate their satisfaction regarding the working conditions in said organisation. In some ESEA methods, the resulting data of such multi-stakeholder surveys is used during the ESEA process. Also in this case, more emphasis should be placed on auditing a potential anomaly. This could be done by making a distinction between multi-stakeholder survey and single stakeholder survey, whereby data from multi-stakeholder survey that is too derived from the average should be recommended. Another possible audit functionality could be to invite the multi-stakeholder survey participants and verify the answers they have filled in.The selection of the respondents could be made on the basis of the same information filled in by the respondent each time. If more than 5 times a survey with the same answer has been filled in, it would be interesting for the auditor to know why the correspondent fills in the same every year. Or if an extremely high/low value is filled in, the participant can be chosen. Finally, a certain approach can be developed to make the analytical features (code in python) created in openESEA work more optimally and less computer power intensive. To summarize, given the fact that this field is undergoing rapid growth and is reliant on research, it is essential to continue facilitating the exchange of information and contribute to the development of ESEA.

9. Conclusion

In this research, we have analysed the audit process of existing ESEA methods and the state of the art on analytical features in computer-assisted auditing tools and techniques (CAATTs). This is done in order to extend openESEA to support the audit with analytical features. This work contributed in the ESEA domain with: Literature study on the origin, terminology, importance and drivers of ESEA, on the audit process of different ESEA methods, a breakdown of audit processes in different fragments, user stories that allow the tool development for this extension, and lastly, a exploration and implementation of several CAATTs in openESEA is performed. To conclude this research, a briefly summary of findings is written.

ESEA has become more and more important to organisations. Therefore these organisations use certain methods to account, audit and assure non financial information. In this research 65 methods are found. All of these methods have a different structure and approach to accounting, reporting, auditing and assuring non-financial data. We have found a collection of 65 methods were all methods have been reviewed and using their accompanying documentation. Finally 10 methods were selected based on existence of an audit component and expert recommendation. For 6 out of 10 methods we have created Process-Deliverable Diagrams (PDDs). These 10 methods were then converted to a super method and a method comparison was conducted. This led to the development of a generic model that supports at least 1 method, namely XES.

Nowadays, almost all audit processes concerning analysis and data extraction are supported by Computer Aided Audit Tools (and 'Techniques'). A variety of techniques have been investigated to apply in openESEA. When studying different techniques, the applicability of the technique was taken into consideration. For example, techniques such as Regression Analysis, which analyzes the relationship of variables and neural networks, which require input layers were not selected. Regardless the fact that Benfords Law is explored more in depth, the decision was made not to include Benford's Law in openESEA yet, since that technique analyses the frequency distribution of the first digit in set of numbers, and this requires historical data. By contrast, three important CAATTs are implemented in openESEA namely, scoring scheme, anomaly detection and visualization methods.

After examining the generic model, and also looking at the process in openESEA, it occured that mainly the activity in the generic model 'Conduct Materiality Risk Assessment' can be improved by CAATTs. Four techniques were chosen to explore more in depth based on the applicability of openESEA and associated capabilities such as data availability, current state of the tool, and the scope of this research related to openESEA. By considering these factors, the decision was made to develop a scoring scheme. The scoring scheme looks at indicators that are decisive in certification type and set to false if they would not be included in the total score. In addition to scoring scheme, a technique called anomaly detection was implemented. This technique filters and recommends the Type 1, extreme values. Given that this method is very powerful for detecting values that are outside the empirical standard, and provides the freedom and opportunity for the auditor to insert their own reasonable ranges this technique is implemented in openESEA. Furthermore, traditional visualization method called boxplot was used. Box plot was chosen given that it fitted well with anomaly detection and reasonable ranges. We have developed an integrated approach to support auditing by introducing CAATTs.

A total of two expert assessments were conducted at different phases of the research during the development. During this validation, the auditing process in openESEA was executed by both experts and the functionalities were tested based on their performance and usability. One of the improvements during the internal validation was on the scoring scheme. This technique did not display the parent of the related indicator. Subsequently, it was found that the presentation of the table to select direct and indirect indicators contained too much unnecessary information (e.g. level, impact) and was in fact confusing and overwhelming. Both comments were enhanced before the external validation. The external evaluation ended with a positive outcome and the experts were satisfied with both the process of the audit and the functionalities based on their usability

Based on an performed evaluation on the functionalities, we were able to validate if we achieve specific goals. This evaluation brought positive results as an outcome and was successfully completed.

References

[1] European parliament. directive 2014/95/eu of the european parliament and of the council of 22 october 2014 amending directive 2013/34/eu as regards disclosure of non-financial and diversity information by certain large undertakings and groups. official journal of the european union, l330(1):1–9, 2014.

[2] K. wiegers. first things first: prioritizing requirements.software development,7(9):48–53, 1999.

[3] Maignan, i., ralston, d. corporate social responsibility in europe and the u.s.: Insights from businesses' self-presentations. j int bus stud 33, 497–514 (2002). https://doi-org.proxy.library.uu.nl/10.1057/palgrave.jibs.8491028.

[4] Ojasoo, m. csr reporting, stakeholder engagement and preventing hypocrisy through ethics audit. j glob entrepr res 6, 14 (2016). https://doi-org.proxy.library.uu.nl/10.1186/s40497-016-0056-9.

[5] Chartered institute of internal auditors. computer assisted audit techniques (caats). 09 2020.

[6] Corporate social responsibility and shared value (esg). *Deloitte*, 2020.

[7] S. A and P. P. *A Review of Financial Accounting Fraud Detection based on Data Mining Techniques*, volume 39. International Journal of Computer Applications (0975 – 8887), 2012.

[8] A. Acquier, J.-P. Gond, and J. Pasquero. Rediscovering howard r. bowen's legacy: The unachieved agenda and continuing relevance of social responsibilities of the businessman. *Business & Society - BUS SOC*, 50:607–646, 12 2011.

[9] C. A. Adams and R. Evans. Accountability, completeness, credibility and the audit expectations gap. *Journal of corporate citizenship*, (14):97–115, 2004.

[10] A. Ahmi and S. Kent. The utilisation of generalized audit software (gas) by external auditors. *Managerial Auditing Journal*, 2013.

[11] L. Albareda and S. Waddock. Networked csr governance: A whole network approach to meta-governance. *Business & Society*, 57(4):636–675, 2018.

[12] E. Ariwa, O. O. Olasanmi, and J. L. Mauri. Green communication and corporate sustainability of computer aided audit techniques and fraud detection. In S.-S. Yeo, Y. Pan, Y. S. Lee, and H. B. Chang, editors, *Computer Science and its Applications*, page 843–862. Springer Netherlands, 2012.

[13] A. Asniarti, I. Muda, et al. The effect of computer assisted audit tools on operational review of information technology audits. In *1st International Conference on Social Sciences and Interdisciplinary Studies (ICSSIS 2018)*, pages 23–27. Atlantis Press, 2019.

[14] B Impact Assessment. The b impact assessment case study, 2022. https://bimpactassessment.net/how-it-works/compare-your-impact.

[15] R. A. Bailey. Design of comparative experiments. *Cambridge University Press*, 2008.

[16] R. D. Banker, H. Chang, and Y.-c. Kao. Impact of information technology on public accounting firm productivity. *Journal of Information Systems*, 16(2):209+, 2002. 209.

[17] F. Benford. The law of anomalous numbers. *Proceedings of the American Philosphical Society*, (78):551–572, 1938.

[18] J. Bierstaker, D. Janvrin, and D. J. Lowe. What factors influence auditors' use of computer-assisted audit techniques? *Advances in Accounting*, 30(1):67–74, 2014.

[19] J. E. Boritz. *Computer Control & Audit Guide.*

[20] H. R. Bowen. *Social responsibilities of the businessman.* University of Iowa Press, 1953.

[21] E. Boyd. History of auditing. *A History of Accounting and Accountants*, 74:703, 1905.

[22] W. C. Boynton and R. N. Johnson. *Modern auditing: Assurance services and the integrity of financial reporting.* John Wiley & Sons, 2005.

[23] M. Bradford, D. Henderson, R. J. Baxter, and P. Navarro. Using generalized audit software to detect material misstatements, control deficiencies and fraud. *Managerial Auditing Journal*, 2020.

[24] R. L. Braun and H. E. Davis. Computer-assisted audit tools and techniques: analysis and perspectives. *Managerial Auditing Journal*, 2003.

[25] S. Brinkkemper. Method engineering: engineering of information systems development methods and tools. *Information and software technology*, 38(4):275–280, 1996.

[26] A. Carroll. Corporate social responsibility: Evolution of a definitional construct. *Business Society*, 38:268–295, 09 1999.

[27] A. Carroll. Corporate social responsibility: Evolution of a definitional construct. *Business & Society*, 38:268–295, 09 1999.

[28] A. B. Carroll and G. W. Beiler. Landmarks in the evolution of the social audit. *Academy of Management Journal*, 18(3):589–599, 1975.

[29] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3), jul 2009.

[30] V. Chiesa, P. Coughlan, and C. A. Voss. Development of a technical innovation audit. *Journal of Product Innovation Management: an international publication of the product development & management association*, 13(2):105–136, 1996.

[31] B. Choi and S. La. The impact of corporate social responsibility (csr) and customer trust on the restoration of loyalty after service failure and recovery. *Journal of Services Marketing*, 2013.

[32] D. Coderre and R. C. M. Police. Global technology audit guide: continuous auditing implications for assurance, monitoring, and risk assessment. *The Institute of Internal Auditors*, pages 1–34, 2005.

[33] M. Cohn. *User stories applied: For agile software development.* Addison-Wesley Professional, 2004.

[34] D. A. Coldwell, J. Billsberry, N. Van Meurs, and P. J. Marsh. The effects of person–organization ethical fit on employee attraction and retention: Towards a testable explanatory model. *Journal of Business Ethics*, 78(4):611–622, 2008.

[35] G. Cosserat and N. Rodda. *Modern auditing.* Wiley, 2009.

[36] Council of European Union. Council regulation (EU) no 95/2014, 2014. `https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32014L0095`.

[37] Council of European Union. Questions and answers: Corporate sustainability reporting directive proposal, 2021. `https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1806`.

[38] E. Creyer and W. T. R. Jr. The impact of corporate behavior on perceived product value. *Marketing Letters*, 7(2):173–185, 1996.

[39] M. B. Curtis and E. A. Payne. An examination of contextual factors and individual characteristics affecting technology implementation decisions in auditing. *International Journal of Accounting Information Systems*, 9(2):104–121, Jun 2008.

[40] M. B. Curtis and E. A. Payne. Modeling voluntary caat utilization decisions in auditing. *Managerial Auditing Journal*, 2014.

[41] P. Dahlin, P. Ekman, J. Röndell, and O. Pesämaa. Exploring the business logic behind csr certifications. *Journal of Business Research*, 112:521–530, 2020.

[42] N. Darnall, I. Seol, and J. Sarkis. Perceived stakeholder influences and organizations' use of environmental audits. *Accounting, Organizations and Society*, 34(2):170–187, 2009.

[43] T. De Magistris, T. Del Giudice, and F. Verneau. The effect of information on willingness to pay for canned tuna fish with different corporate social responsibility (csr) certification: a pilot study. *Journal of Consumer Affairs*, 49(2):457–471, 2015.

[44] R. Debreceny, S.-L. Lee, W. Neo, and J. S. Toh. Employing generalized audit software in the financial services sector: Challenges and opportunities. *Managerial Auditing Journal*, 20(6):605–618, 2005.

[45] C. Deegan, B. J. Cooper, and M. Shelly. An investigation of tbl report assurance statements: Uk and european evidence. *Managerial Auditing Journal*, 2006.

[46] M. A. Delmas and V. C. Burbano. The drivers of greenwashing. *California management review*, 54(1):64–87, 2011.

[47] T. M. J. F. Derikx. Designing a domain-specific language for infographics. master's thesis. Utrecht Universiteit, 2020.

[48] C. E. Dictionary. Accounting, noun. In *Cambridge English Dictionary*. Cambridge University Press, 2021.

[49] C. E. Dictionary. Audit, noun. In *Cambridge English Dictionary*. Cambridge University Press, 2021.

[50] C. E. Dictionary. Auditor's report, noun. In *Cambridge English Dictionary*. Cambridge University Press, 2021.

[51] R. G. Eccles, L.-E. Lee, and J. C. Stroehle. The social origins of esg: An analysis of innovest and kld. *Organization & Environment*, pages 575–596, 2020.

[52] S. G. Eick and D. E. Fyock. Visualizing corporate data. *AT&T technical journal*, 75(1):74–86, 1996.

[53] M. J. Epstein and A. R. Buhovac. *Making sustainability work: Best practices in managing and measuring corporate social, environmental, and economic impacts*. Berrett-Koehler Publishers, 2014.

[54] S. España and S. Brinkkemper. Responsible software: A research agenda to help enterprises become more sustainable. In *ICT for Sustainability 2016*, pages 141–150. Atlantis Press, 2016.

[55] S. Espana and S. Brinkkemper. Responsible software: A research agenda to help enterprises become more sustainable. In *Proceedings of ICT for Sustainability 2016*, pages 141–150. Atlantis Press, 2016/08.

[56] S. España, N. Bik, and S. Overbeek. Model-driven engineering support for social and environmental accounting. pages 1–12. 13th International Conference on Research Challenges in Information Science (RCIS), 2019.

[57] European Commission. Proposal for a directive of the european parliament and of the council amending directive 2013/34/eu, directive 2004/109/ec, directive 2006/43/ec and regulation (eu) no 537/2014, 2021. `https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021PC0189`.

[58] European Council of the European Union. Information assurance, 2021. https://www.consilium.europa.eu/en/general-secretariat/corporate-policies/classified-information/information-assurance/.

[59] F.-F. d. E. C. Européens. Fee call for action: Assurance for sustainability. *Brussels: FEE*, 2004.

[60] E. Feroz, T. Kwon, V. Pastena, and K. Park. The efficacy of red flags in predicting the sec's targets: An artificial neural networks approach. *Intelligent Systems in Accounting Finance Management*, 29:145–157, 09 2000.

[61] A. Ferrari, P. Spoletini, M. Bano, and D. Zowghi. Learning requirements elicitation interviews with role-playing, self-assessment and peer-review. pages 28–39, 09 2019.

[62] J. Flower. The international integrated reporting council: A story of failure. *Critical Perspectives on Accounting*, 27:1–17, 2015.

[63] R. Foorthuis. *A Typology of Data Anomalies.* 2018.

[64] R. Foorthuis. SECODA: segmentation- and combination-based detection of anomalies. *CoRR*, abs/2008.06869, 2020.

[65] M. Friedman. The social responsibility of business is to increase its profits. *Heidelberg: Springer Berlin Heidelberg*, page 173–178, 1970.

[66] M. Galarnyk. Understanding boxplots. 2018.

[67] J. Galbreath and P. Shum. Do customer satisfaction and reputation mediate the csr–fp link? evidence from australia. *Australian journal of management*, 37(2):211–229, 2012.

[68] Y. Ganesan, Y. W. Hwa, A. H. Jaaffar, and F. Hashim. Corporate governance and sustainability reporting practices: the moderating role of internal audit function. *Global Business and Management Research*, 9(4s):159–179, 2017.

[69] E. Garriga and D. M. e. Corporate social responsibility theories:mapping the territory. *Journal of business ethics*, 53:51–71, 2004.

[70] R. Ghani, K. Rosli, N. A. Ismail, and S. Z. Saidin. Application of computer-assisted audit tools and techniques (caatts) in audit firms. *Journal of Advanced Research in Business and Management Studies*, 9(1):67–74, 2017.

[71] D. C. GmbH. Sustainability reporting: Third-party assurance. 2016.

[72] J. Graafland and B. Van de Ven. Strategic and moral motivation for corporate social responsibility. *Journal of Corporate Citizenship*, (22):111–123, 2006.

[73] R. Gray. Current developments and trends in social and environmental auditing, reporting and attestation: a review and comment. *International journal of auditing*, 4(3):247–268, 2000.

[74] R. Gray. Social, environmental and sustainability reporting and organisational value creation?: Whose value? whose creation? *Accounting, Auditing Accountability Journal*, 19:793–819, 11 2006.

[75] R. Gray and J. Bebbington. Environmental accounting, managerialism and sustainability: Is the planet safe in the hands of business and accounting? advances in environmental accounting management. 1:1–44, 2000.

[76] M. e. a. Hasan. The different types of assurance services and levels of assurance provided. *International Journal of Auditing*, 9(2):91–102, 2005.

[77] O. Hinkelmann, Klaus; Kempthorne. Design and analysis of experiments. *Cambridge University Press*, Volume I: Introduction to Experimental Design, 2008.

[78] D. E. Hirst and L. Koonce. Audit analytical procedures: A field investigation. *Contemporary Accounting Research*, 13(2):457–486, 1996.

[79] K. Hodge, N. Subramaniam, and J. Stewart. Assurance of sustainability reports: Impact on report users' confidence and perceptions of information credibility. *Australian accounting review*, 19(3):178–194, 2009.

[80] S.-M. Huang, Y.-C. Hung, and H.-H. Tsao. Examining the determinants of computer-assisted audit techniques acceptance from internal auditors' viewpoints. 4(4):377 – 392, 2008.

[81] IAASB. Isae 3000 (revised), assurance engagements other than audits or reviews of historical financial information. 2013.

[82] S. O. Idowu, N. Capaldi, L. Zu, and A. D. Gupta. *Encyclopedia of corporate social responsibility*, volume 21. Springer New York, 2013.

[83] International Auditing and Assurance Standards Board. Assurance engagements completed, 2021.
`https://www.iaasb.org/projects/assurance-engagements-completed`.

[84] ISACA. Generalized audit software: Effective and efficient tool for today's it audits. page 3, 2006.

[85] N. Ivers, G. Jamtvedt, S. Flottorp, J. M. Young, J. Odgaard-Jensen, S. D. French, M. A. O'Brien, M. Johansen, J. Grimshaw, and A. D. Oxman. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane database of systematic reviews*, (6), 2012.

[86] A. J. Izenman. Modern multivariate statistical techniques: Regression, classification, and manifold learning. 2008.

[87] K. S. Jahdi and G. Acikdilli. Marketing communications and corporate social responsibility (csr): marriage of convenience or shotgun wedding? *Journal of business ethics*, 88(1):103–113, 2009.

[88] D. Janvrin. Auditor acceptance of computer-assisted audit techniques. Jun 2008.

[89] G. Johnson, A. Varnon, and S. Zhang. Using benford's law and mean absolute deviation to distinguish bankrupt fraudulent companies and healthy currently operating companies. *Southeast Missouri State University*, (78):114–142, 2014.

[90] C. K and Y. J. *A review of Fraud Detection Techniques: Credit Card*, volume 45. International Journal of Computer Applications (0975 – 8887), 2012.

[91] C. H. Kim and H. Scullion. The effect of corporate social responsibility (csr) on employee motivation: A cross-national study. *Poznan university of economics review*, 13(2), 2013.

[92] H.-J. Kim, M. Mannino, and R. J. Nieschwietz. Information technology acceptance in the internal audit profession: Impact of technology features and complexity. *International Journal of Accounting Information Systems*, 10(4):214–228, Dec 2009.

[93] S. Kim. The process model of corporate social responsibility (csr) communication: Csr communication and its relationship with consumers' csr knowledge, trust, and corporate reputation perception. *Journal of business ethics*, 154(4):1143–1159, 2019.

[94] S. Kim and M. Ferguson. Public expectations of csr communication: What and how to communicate csr. *Public Relations Journal*, 8(3):1–22, 2014.

[95] B. Krisviandik. *The ASQ Auditing Handbook*. Number 3. 2005.

[96] H. E. Kruiper. Model-driven engineering of ethical, social and environmental accounting stakeholder surveys. master's thesis. Utrecht Universiteit, 2020.

[97] G. P. Lantos. The boundaries of strategic corporate social responsibility. *Journal of consumer marketing*, 2001.

[98] T. Laudal. Drivers and barriers of csr and the size and internationalization of firms. *Social Responsibility Journal*, 2011.

[99] L. Lee and L.-F. Chen. Boosting employee retention through csr: A configurational analysis. *Corporate Social Responsibility and Environmental Management*, 25(5):948–960, 2018.

[100] D. J. Lowe, J. L. Bierstaker, D. J. Janvrin, and J. G. Jenkins. Information technology in an audit context: Have the big 4 lost their advantage? *Journal of information systems*, 32(1):87–107, 2018.

[101] X. Luo and C. B. Bhattacharya. Corporate social responsibility, customer satisfaction, and market value. *Journal of marketing*, 70(4):1–18, 2006.

[102] N. Mahzan and A. Lymer. Adoption of computer assisted audit tools and techniques (caatts) by internal auditors: Current issues in the uk. page 47.

[103] N. Mahzan and A. Lymer. Examining the adoption of computer-assisted audit tools and techniques. *Managerial Auditing Journal*, 2014.

[104] S. Maijoor, A. Vanstraelen, and T. Mock. The determinationa and communication of levels of assurance other than high. IFAC, 2002.

[105] S. Maijoor, A. Vanstraelen, and T. Mock. Sustainability reporting and assurance: A historical analysis on a world-wide phenomenon. volume 120, pages 1–11. Journal of Business Ethics, 2014.

[106] J. Maltby. Environmental audit: theory and practices. *Managerial auditing journal*, 1995.

[107] P. Martínez and I. R. Del Bosque. Csr and customer loyalty: The roles of trust, customer identification with the company and satisfaction. *International Journal of Hospitality Management*, 35:89–99, 2013.

[108] D. Matthews. A history of auditing: The changing audit process in britain from the nineteenth century to the present day. pages 1–208. Routledge, 2006.

[109] M. P. Miles and L. S. Munilla. The potential impact of social accountability certification on marketing: A short note. *Journal of Business Ethics*, 50(1):1–11, 2004.

[110] M. Minnis. The value of financial statement verification in debt financing: Evidence from private us firms. *Journal of accounting research*, 49(2):457–506, 2011.

[111] T. J. Mock, C. Strohm, and K. M. Swartz. An examination of worldwide assured sustainability reporting. 17(41):67–77, 2007.

[112] Z. Mohamed, M. Zain, and W. F. Wan Yusoff. Internal audit attributes and external audit's reliance on internal audit: Implications for audit fees. *International Journal of Auditing*, 16:268–285, 11 2012.

[113] L. A. Mohr, D. J. Webb, and K. E. Harris. Do consumers expect companies to be socially responsible? the impact of corporate social responsibility on buying behavior. *Journal of Consumer Affairs*, 35(1):45–72, 2001. `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1745-6606.2001.tb00102.x`.

[114] U. Nation. Who cares wins: Connecting financial markets to a changing world. *The Global Compact*, page 1–44, 2004. `https://www.unglobalcompact.org/docs/issues_doc/Financial_markets/who_cares_who_wins.pdf`.

[115] M. Nelson. How can climate change disclosures protect reputation and value? *Ernst & Young*, 2020.

[116] M. J. Nigrini. *Benford's Law: Applications for forensic accounting, auditing, and fraud detection*, volume 586. John Wiley & Sons, 2012.

[117] M. J. Nigrini and P. D. Drake. Computer assisted analytical procedures using benford's law. *Journal of accounting education*, pages 127–146, 2000.

[118] M. J. Nigrini and L. J. Mittermaier. The use of benford's law as an aid in analytical procedures. *Auditing*, 16(2):52, 1997.

[119] I. C. of Commerce. *An ICC guide to effective environmental auditing.* ICC Publ., 1991.

[120] M. Ojasoo. Csr reporting, stakeholder engagement and preventing hypocrisy through ethics audit. *Journal of Global Entrepreneurship Research*, 6(1):1–14, 2016.

[121] B. O'Dwyer, D. Owen, and J. Unerman. Seeking legitimacy for new assurance forms: The case of assurance on sustainability reporting. *Accounting, Organizations and Society*, 36(1):31–52, 2011.

[122] B. O'Dwyer and D. L. Owen. Assurance statement practice in environmental, social and sustainability reporting: a critical evaluation. 37(2):205–229, 2011.

[123] R. S. Pediredla, V. Ravi, G. Rao, and I. Bose. Detection of financial statement fraud and feature selection using data mining techniques. *Decision Support Systems*, 50:491–500, 01 2011.

[124] I. Pedrosa and C. Costa. Computer assisted audit tools and techniques in real world: Caatt's applications and approaches in context. *t, International Journal of Computer Information Systems and Industrial Management Applications*, 4:161–168, 12 2012.

[125] P. Perego and A. Kolk. Multinationals' accountability on sustainability: The evolution of third-party assurance of sustainability reports. *Journal of Business Ethics*, 110, 10 2012.

[126] A. Pérez and I. R. del Bosque. Customer csr expectations in the banking industry. *International Journal of Bank Marketing*, 2014.

[127] F. Perrini. The practitioner's perspective on non-financial reporting. page 73–103. California Management Review, 2012.

[128] S. Phillips, V. V. Thai, and Z. Halim. Airline value chain capabilities and csr performance: the connection between csr leadership and csr culture with csr performance, customer satisfaction and financial performance. *The Asian Journal of Shipping and Logistics*, 35(1):30–40, 2019.

[129] M. A. F. Pimentel, D. A. Clifton, L. A. Clifton, and L. Tarassenko. A review of novelty detection. *Signal Process.*, 99:215–249, 2014.

[130] B. Porter, J. Simon, and D. Hatherly. *Principles of external auditing*. John Wiley & Sons, 2003.

[131] V. Ramautar. The state of the art and practice on social and environmental accounting methods and tools. Master's thesis, Utrecht University, 2019.

[132] V. Ramautar. The state of the art and practice on social and environmental accounting methods and tools. master's thesis. Utrecht Universiteit, 2019.

[133] W. Randy Evans and W. D. Davis. An examination of perceived corporate citizenship, job applicant attraction, and csr work role definition. *Business & Society*, 50(3):456–480, 2011.

[134] P. Roselle. The evolution of integrating esg analysis into wealth management decisions. page 75–79. Journal of Applied Corporate Finance, 2016.

[135] M. Salehi. Reasonableness of audit expectation gap: Possible approach to reducing. *ICFAI Journal of Audit Practice*, 4(3), 2007.

[136] C. Seow, R. Hillary, S. S. Gao, and J. J. Zhang. Stakeholder engagement, social auditing and corporate sustainability. *Business process management journal*, 2006.

[137] E.-G. Siew, K. Rosli, and P. H. Yeow. Organizational and environmental influences in the adoption of computer-assisted audit tools and techniques (caatts) by audit firms in malaysia. *International Journal of Accounting Information Systems*, 36:100445, 2020.

[138] R. Simnett, A. Vanstraelen, and W. Chua. Assurance on sustainability reports: An international comparison. *Accounting Review*, 84, 10 2007.

[139] R. Simnett, A. Vanstraelen, and W. F. Chua. Assurance on sustainability reports: An international comparison. *The accounting review*, 84(3):937–967, 2009.

[140] B. A. Sparks, H. E. Perkins, and R. Buckley. Online travel reviews as persuasive communication: The effects of content type, source, and certification logos on consumer behavior. *Tourism Management*, 39:1–9, 2013.

[141] C. Spathis, M. Doumpos, and C. Zopounidis. Using client performance measures to identify pre-engagement factors associated with qualified audit reports in greece. *The International Journal of Accounting*, 38:267–284, 09 2003.

[142] K. Stephens and M. T. Roszak. *A study of the role and benefits of third party auditing in Quality Management Systems*, volume 43. Journal of Achievements in Materials and Manufacturing Engineering, 2010.

[143] M. H. Taylor and D. T. Simon. Determinants of audit fees: the importance of litigation, disclosure, and regulatory burdens in audit engagements in 20 countries. *The International Journal of Accounting*, 34(3):375–388, 1999.

[144] L. Teck-Heang and A. M. Ali. The evolution of auditing: An analysis of the historical development. *Journal of Modern Accounting and Auditing*, 4(12):1, 2008.

[145] O. Teofilovski. Thomson reuters esg scores. page 12. Utrecht Universiteit, 2017.

[146] R. Thompson, C. Higgins, and J. Howell. Personal computing: Toward a conceptual model of utilization. mis quarterly. pages 125–143, 1991.

[147] S. Turley and M. Cooper. *Auditing in the United Kingdom: A study of development in the audit methodologies of large accounting firms*. Prentice Hall, 1991.

[148] I. van de Weerd and S. Brinkkemper. Meta-modeling for situational analysis and design methods. In *Handbook of research on modern systems analysis and design technologies and applications*, pages 35–54. IGI Global, 2009.

[149] I. van de Weerd, S. de Weerd, and S. Brinkkemper. Developing a reference method for game production by method comparison. In *Working Conference on Method Engineering*, pages 313–327. Springer, 2007.

[150] T. van Dijk. Validating your environmental efforts:designing audit functionality for auniversal esea tool. Utrecht University, 2022.

[151] S. van Nijen. An analysis of the social impact assessment family of methods. Utrecht University, 2020.

[152] V. Venkatesh. A model of the antecedents of perceived peou: Development and test. decision science. pages 415–450, 1996.

[153] C. R. W. Descriptive statistics for summarising data. (1):61–139, 2020.

[154] P. Wallage. Assurance on sustainability reporting: An auditor's view. 19, 2000.

[155] S. Wang. A comprehensive survey of data mining-based accounting-fraud detection research. In *2010 International Conference on Intelligent Computation Technology and Automation*, volume 1, pages 50–53. IEEE, 2010.

[156] S. Wang. A comprehensive survey of data mining-based accounting-fraud detection research. In *2010 International Conference on Intelligent Computation Technology and Automation*, volume 1, pages 50–53, 2010.

[157] R. J. Wieringa. *Design science methodology for information systems and software engineering*. Springer, 2014.

[158] D. Williamson, R. Parker, and J. Kendrick. The box plot: A simple visual method to interpret data. *Annals of internal medicine*, 110:916–21, 07 1989.

[159] W. Zhou and G. Kapoor. Detecting evolutionary financial statement fraud. *Decision Support Systems*, 50:570–575, 02 2011.

A. PDDs

A.1. EMAS

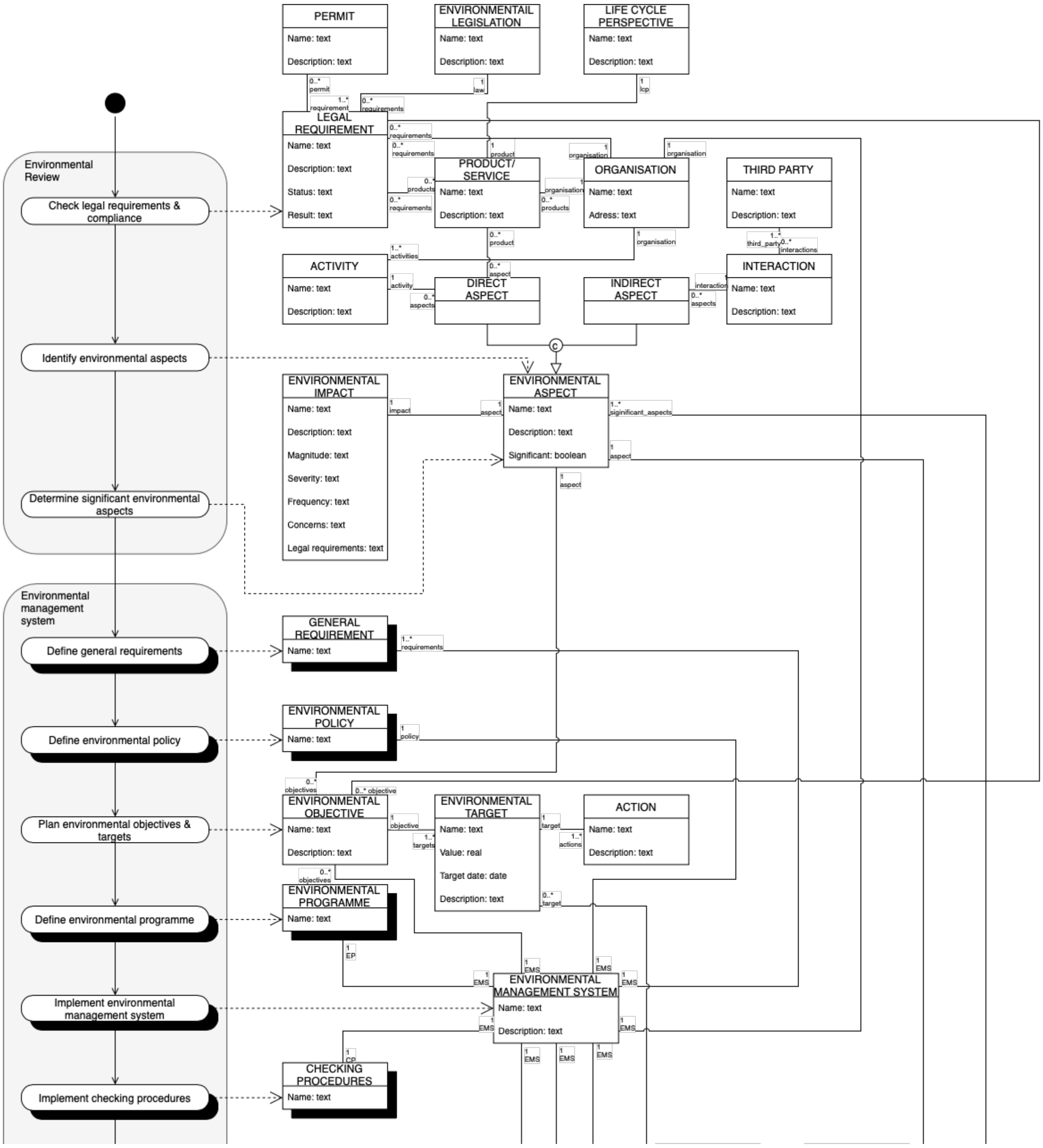

Figure 14: EMAS PDD Part I

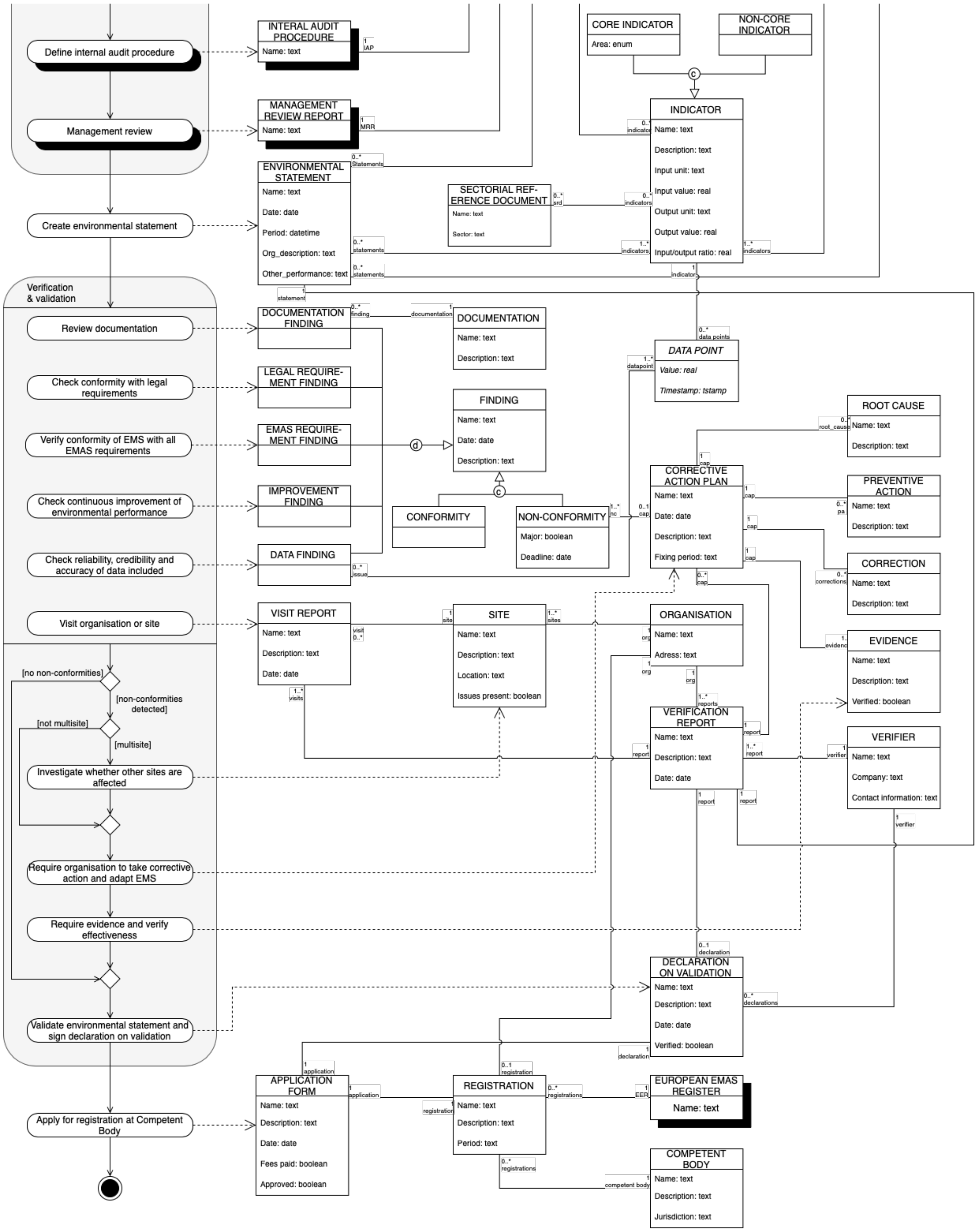

Figure 15: EMAS PDD Part II

A.2. ISAE 3000

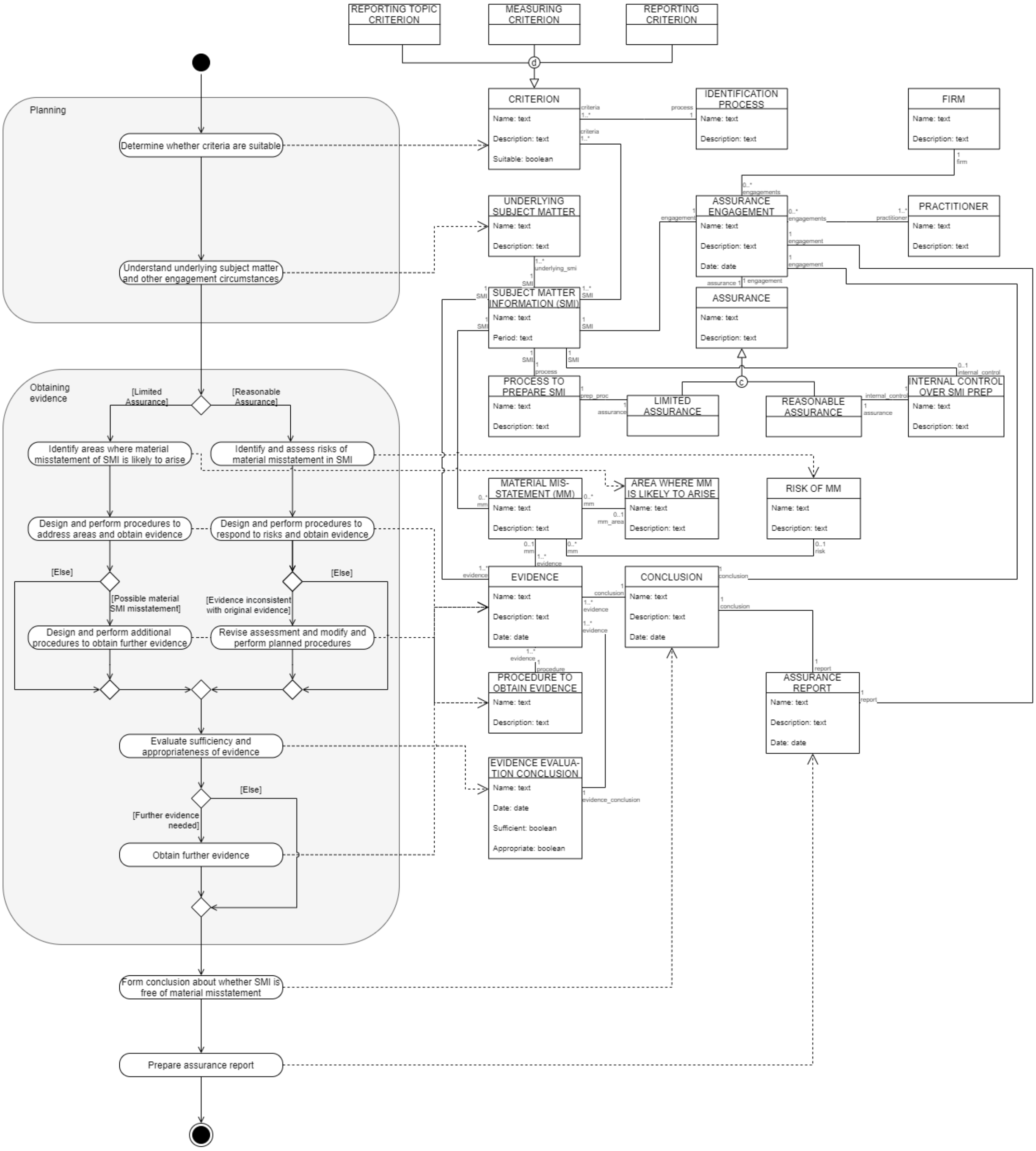

Figure 16: ISAE 3000 PDD

A.3. Green House Gases Protocol

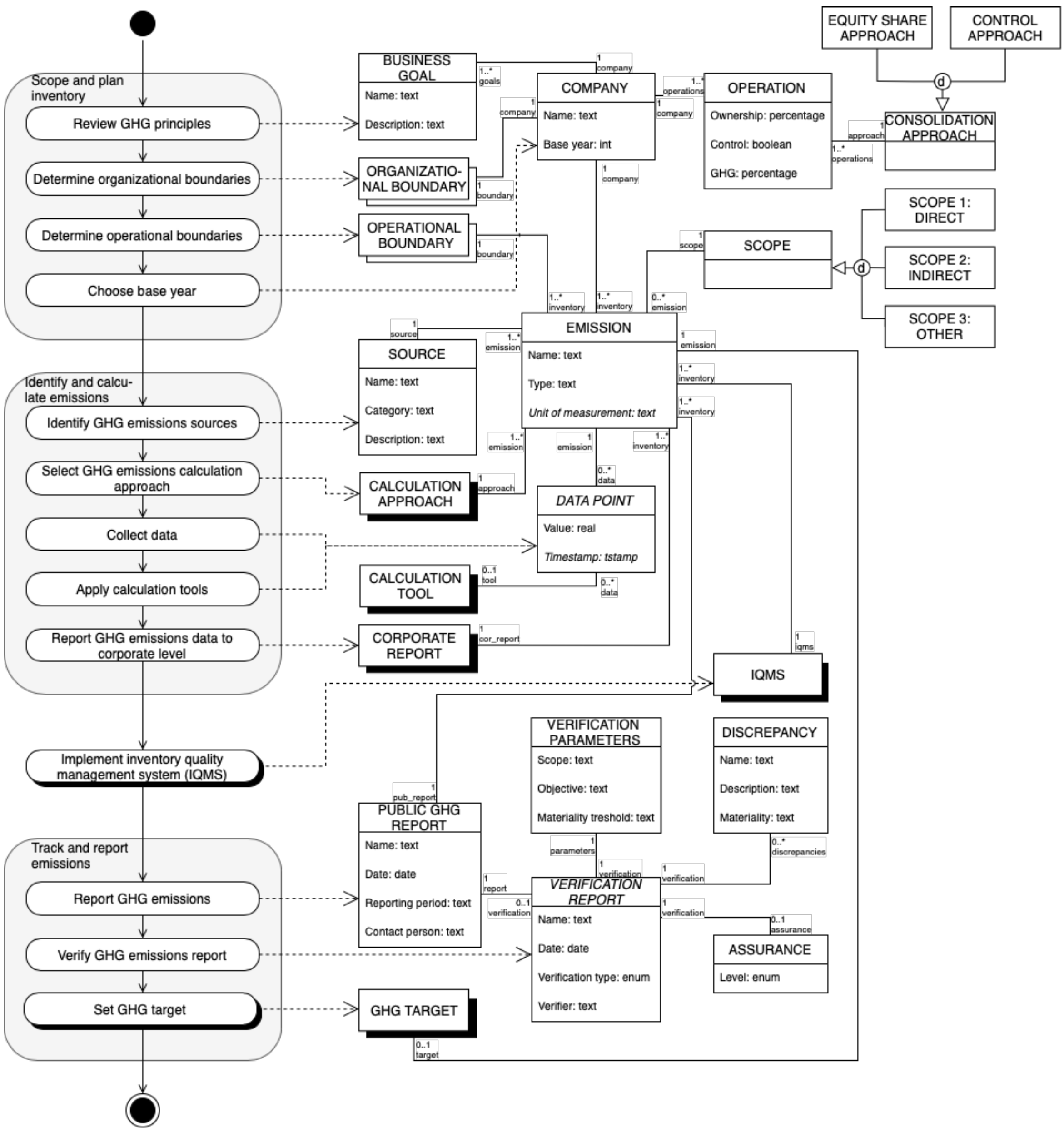

Figure 17: GHG PDD

A.4. Natural Capital Protocol

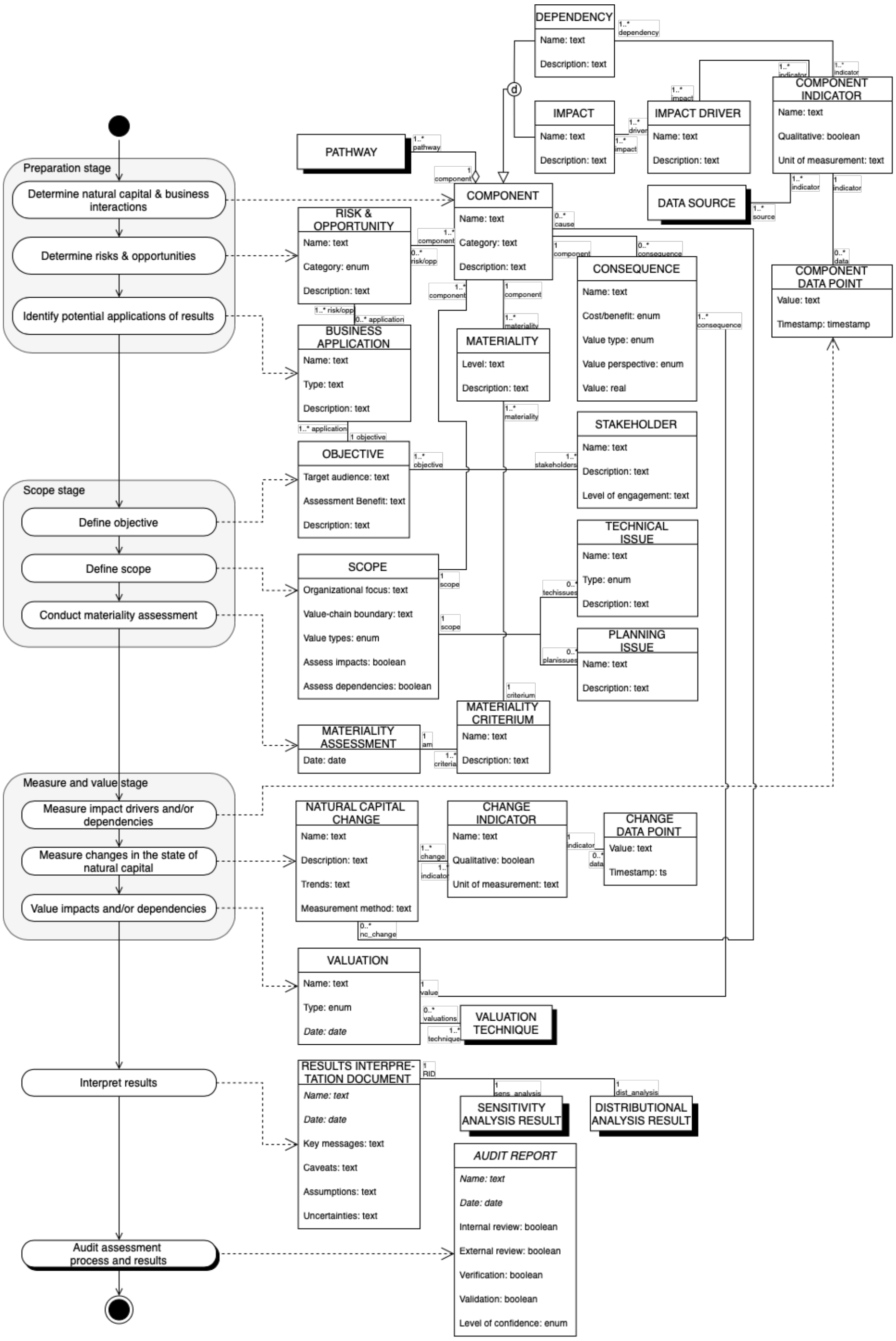

Figure 18: NCP PDD

A.5. B Impact Assessment

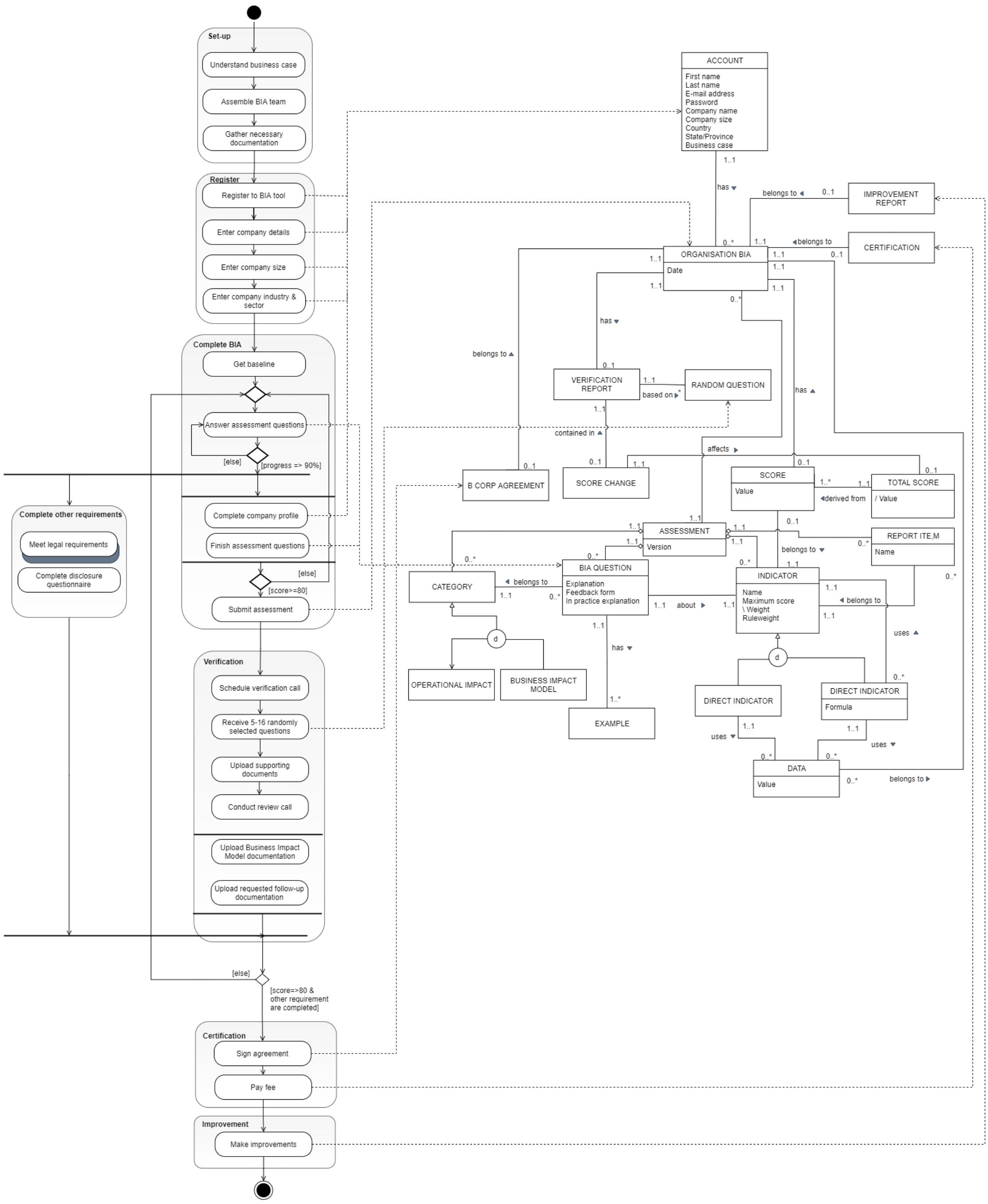

Figure 19: B Impact Assessment PDD

A.6. ISO 14011

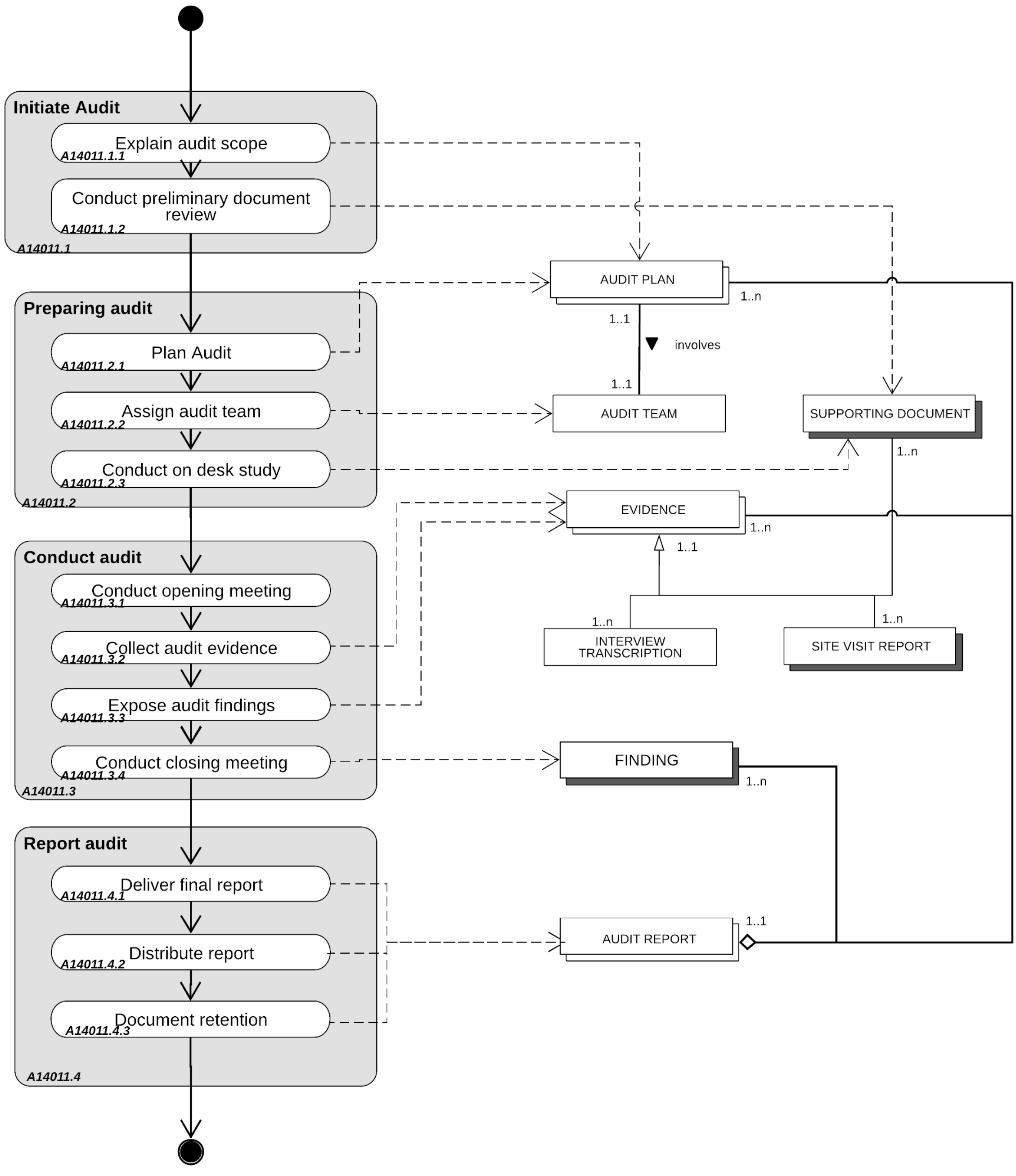

Figure 20: ISO 14011 PDD

A.7. AA1000

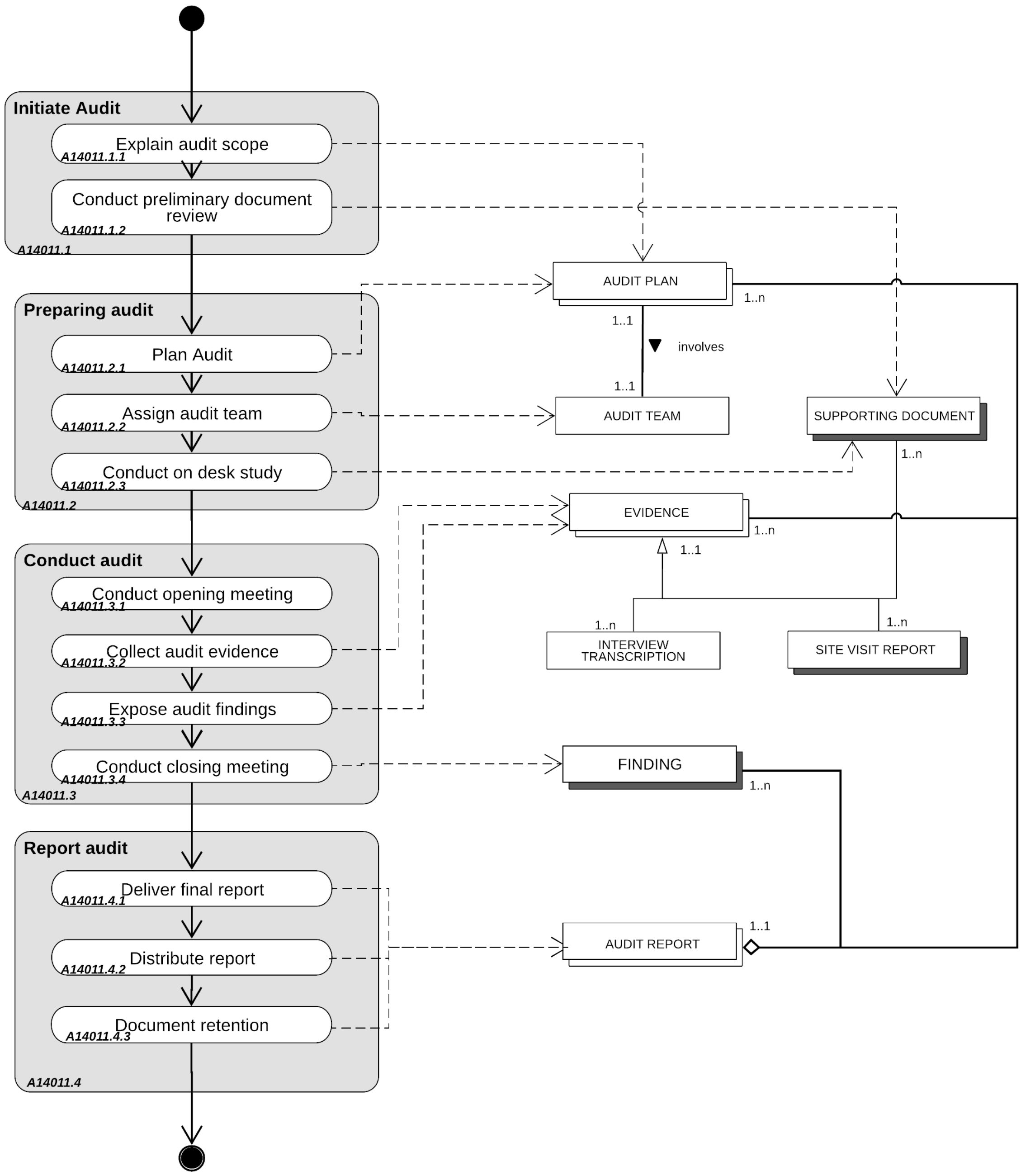

Figure 21: AA1000 PDD

A.8. XES Social Balance audit

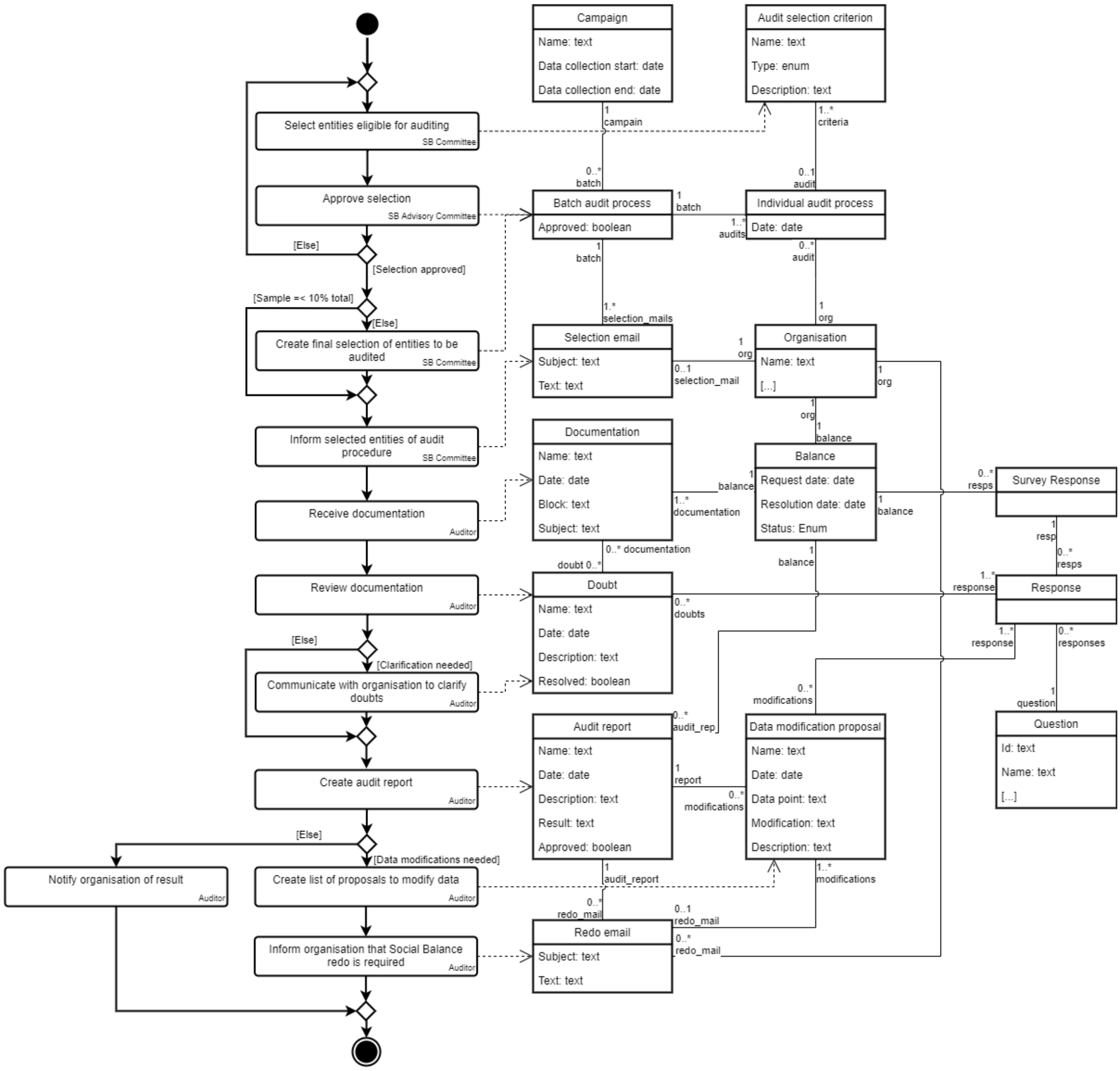

Figure 22: XES Social Balance PDD

A.9. FLA-SCI

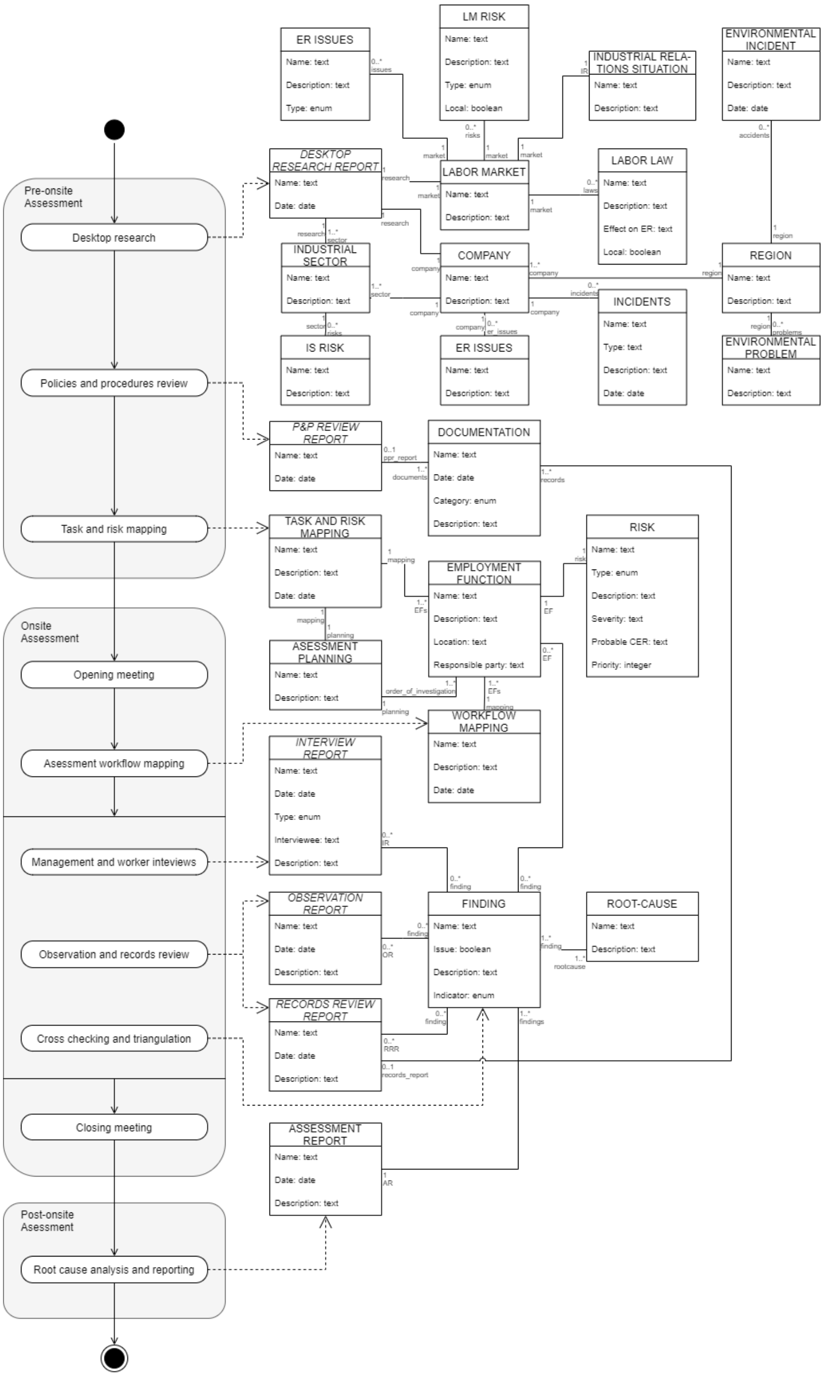

Figure 23: FLA-SCI PDD

A.10. GRI

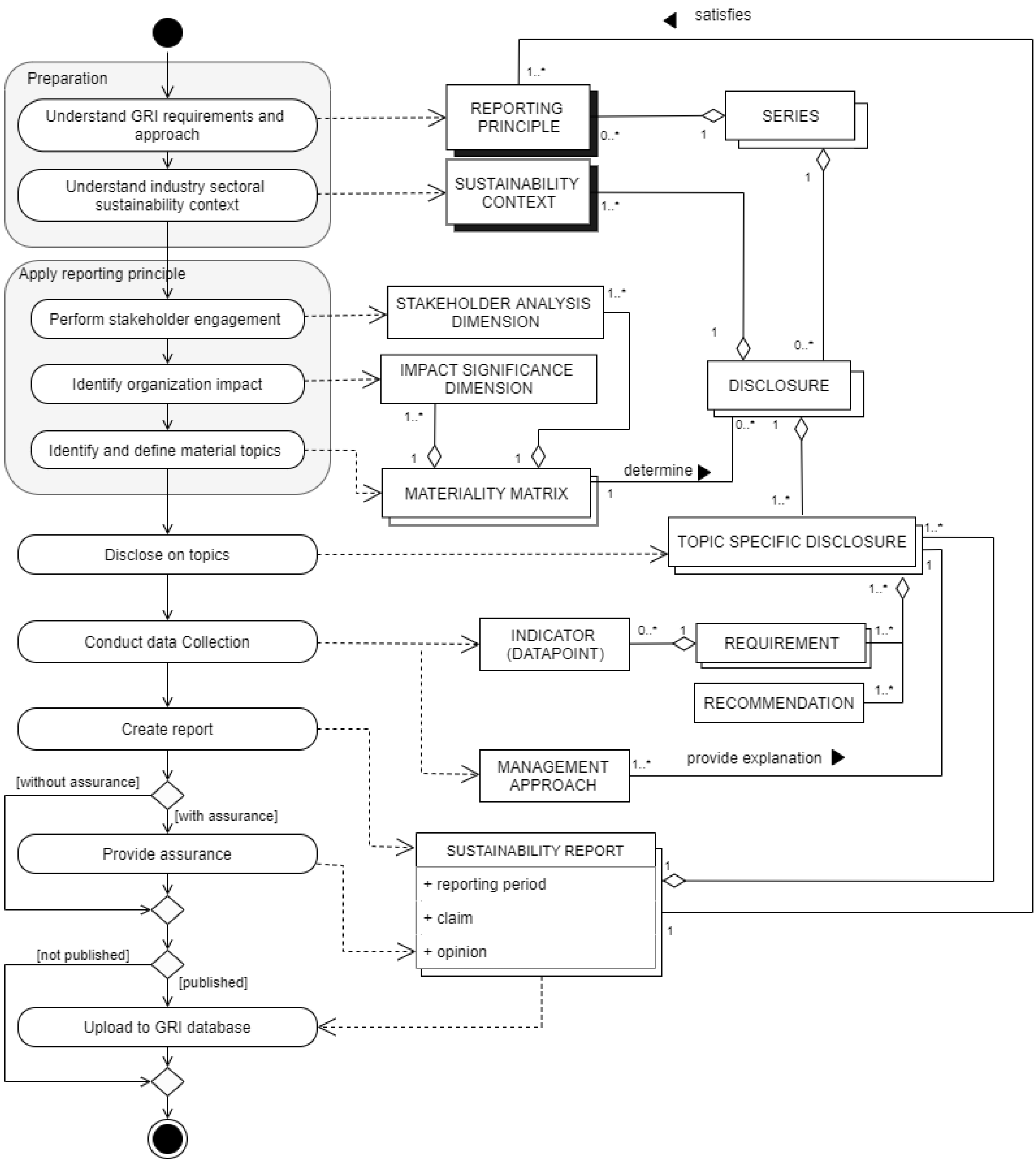

Figure 24: GRI PDD

B. PDD Validation Interview Protocol

First of all, thank you for making time for this interview. Before we get started we have to ask for your permission to record this meeting.

Question 1: Is it okay if we record our meeting?

- Yes: You may start the recording
- No: Take notes during the meeting

Note that the recording has started.

Introduction to the interview
This interview is part of our thesis which focuses on audit functionalities of what we call 'ethical social environmental accounting.' As part of our thesis we have modelled several ESEA methods. We have created so-called Process Delivery Diagrams that provide a slightly abstract overview of the approach of these methods. These models are based on method documentation or, if no documentation is available, we have used the online or offline tools that have been developed to support practitioners of the method to distill the approach of the method. Through these models we try to get a better understanding of how your method is performed. The goal of this interview is to ensure that the model reflects the process as intended by the developers and/or interpreted by consultants and practitioners. This is mainly done by comparing your impressions of the method to our model. With impressions we mean the underlying intentions of the method and how it should be executed.

Question 1: Do you have any questions before we get started?

- Yes: Answer questions
- No: Proceed to **Question 3**

General Questions
Question 3: What is the name of your job function and what does it entail?
Question 4: How does your job relate to the usage of the [method/tool]?

- What is your relationship to the [method/tool] and how did you get to know about this [method/tool]?
- Since how long are you familiar with the [method/tool]?

PDD explanation
Before we jump into the validation it is relevant to provide you with a bit of context on how we modelled the [method] and how to interpret it. We have built a so-called process-deliverable diagram (PDD) of the [method]. This diagram consists of two parts, a process part and a deliverable part. We will explain in short what it entails. The process part of the model describes the activities of the [method], the relations among them, the order in which they should be performed and which actors are responsible for the completion of these activities. Show an example of the process part of stage 1 of the [method]. The deliverable part of the model is related to the process part of the [method]. A deliverable is nothing more than an output resulting from an activity. Show an example of a deliverable resulting from an activity of the [method].

Process validation
Very well. We can start the validation process now. We will first go through the process part of the model.
First, we show the process part of the PDD and read out the names of the stages

Question 5: Do you recognize all the stages of the method that we define in the model?

- Yes: Go to the next sub-question
- No: Which stages are you not familiar with?

Question 6: Are there perhaps any parts of the impact measurement process that you as a [consultant prescribe to your clients/practitioners to apply] that are not mentioned in the method documentation or covered in the tool?

- Yes: Which stages would you define on top of the ones that we mention in our model?
- No: Go to **Question 7**

Within the stages we have identified activities. The goal of this question is to determine if the activities in the PDD correspond with the activities defined by the [tool/documentation]. This is done by showing the activity in the [tool/documentation] and comparing it with the activity in the PDD. We repeat the following questions for each stage in the process part.

Question 7: Does the activity in the PDD reflect the corresponding activity in the [documentation/tool]?

- Yes: Go the next activity
- No: What is different and how to change this?

We repeat Question 7 until all activities in this stage are discussed

Question 8: Are there any activities missing in this stage that should be there according to [company/de- veloper name/interviewee]?

- Yes: Which activity is missing and how do you describe the activity?
- No: Go to **Question 9**

Question 9: Are the activities in this stage in the order that corresponds with the [tool/documentation]?

- Yes: Go to **Question 10**
- No: Which activity does not correspond and why?

We repeat these processes until all activities are discussed.

Question 10: Do you have any other comments on this stage of the [method]?

- Yes: Discuss comments
- No: Continue to Deliverable Validation

Deliverable validation

Now that we have validated the process part, let's take a look at the deliverables resulting from the activities. Just as a reminder, deliverables are data and documents that result from activities. We will go over the deliverables stage by stage. For each phase, we discuss the deliverables of the activities. The following questions will be asked for each activity of each phase.

Question 11: Does the deliverable from [activity] correspond with the [documentation/tool]?

- Yes: Go to the next deliverable
- No: Why not and what should be changed?

Question 12: Are there any critical deliverables missing in this stage?

- Yes: Which deliverable is missing and why?
- No: Continue.

Ending of interview
The person being interview is thanked for the time spend on the interview. Further possible collaborations are being discussed and if there is time left a demo of openESEA is showed.

C. User Roles

User roles

ID	Name	Description	Type	Includes
RA	Root administrator	The user who manages the installation of the software. Can do virtually everything in the backend: manage networks, organisations, engineer methods, etc.	Concrete	ME
ME	Method engineer	The user who specifies an Impact Measurement method	Abstract	
NA	Network administrator	The user who manages a network. Can manage a network, create organisations within that network, create aggregated benchmark reports, export data of the whole network, etc.	Concrete	ME
OA	Organisation administrator	Can engineer methods, manage their organisation, etc.	Concrete	ME
IM	IM accountant/practitioner	Can execute Impact Measurement (IM) methods, create Theory of Changes, activate and send surveys, validate data and edit it if necessary, close surveys, produce reports and send them to stakeholders, produce reports/infographics and send them to stakeholders, export data, etc.	Concrete	
SR	Survey respondent	Can respond to surveys that are sent to them, enter requested data, download received reports.	Concrete	
AU	Auditor	Can review (but not edit) data to check the accuracy and materiality, sign electronically to attest the result. (Not to be implemented yet, but in later sprints)	Concrete	

Figure 25: User roles of openESEA

ID	Activity	Sub-activity	Description
A	Initiate audit		
A.1		Determine audit scope	Determining the AUDIT SCOPE. The audit scope is often agreed upon by the AUDITOR and the REPORTING ORGANISATION in consultation with each other.
A.2		Conduct preliminary documentation review	The AUDITOR reviews preliminary documentation on the REPORTING ORGANISATION, to get a better view of the organisation and it's undertakings. When a NON-FINANCIAL REPORT is audited, the auditor also reviews this to get a better understanding of the audit engagement.
A.3		Select organisations eligible for auditing	Based on the AUDIT SELECTION CRITERIA, organisations that are eligible for auditing are selected. This results in the BATCH AUDIT PROCESS.
A.4		Approve selection of organisations	The selection of REPORTING ORGANISATIONS (the BATCH AUDIT PROCESS) is approved.
A.5		Create final selection of organisations to be audited	When the BATCH AUDIT PROCESS comprises too many REPORTING ORGANISATIONS, a selection of these organisations is made. This results in a finalized BATCH AUDIT PROCESS.
A.6		Inform selected organisations of audit	The selected REPORTING ORGANISATIONS are informed that they will be audited.
B	Prepare audit		
B.1		Determine procedures to be used during audit	The procedures that are to be used during the audit are determined. This might include, for example, interviews, site visits and documentation requests.
B.2		Determine audit timeline	The timeline of the audit is determined. This might include a start date, end date or other deadlines for the REPORTING ORGANISATION and/or AUDITOR.
B.3		Identify audit team members	Audit team members are identified. In many audit processes, multiple AUDITORS work together on the AUDIT ENGAGEMENT.
B.4		Assign actions to audit team	Auditing actions are assigned to the audit team.
B.5		Identify functions within organisation responsible for documentation	The functions in the REPORTING ORGANISATION are identified that are responsible for specific parts of the documentation.
B.6		Understand underlying subject matter	The AUDITOR reviews the UNDERLYING SUBJECT MATTER.
B.7		Conduct risk of material misstatement assessment	For all data that is audited, the risk is assessed that said data contains material misstatements. The risk assessment might, for example, include variance analysis, historical comparison and comparisons with peers.
B.8		Schedule audit call	A call is scheduled for the audit.
C	Conduct audit		
C.1		Receive randomly selected questions	Random questions are selected for which the REPORTING ORGANISATION has provided answers. The responses of the selected questions will be audited. For these responses, the REPORTING ORGANISATION will have to provide documentation.
C.2		Conduct opening meeting	An opening meeting is conducted for the AUDITOR and the REPORTING ORGANISATION. The main goal is to lay the foundation for the working relationship with the REPORTING ORGANISATION.
C.3.a		**Elicit information:** Request documentation	DOCUMENTATION is requested for the data that is audited.
C.3.b		**Elicit information:** Visit organisation	The REPORTING ORGANISATION is visited by the AUDITOR to elicit information that supports the audited data.
C.3.c		**Elicit Information:** Conduct interviews	Employees of the REPORTING ORGANISATION are interviewed by the AUDITOR to elicit information that supports the audited data.
C.3.d		**Elicit information:** Cross checking and triangulation	The audit team comes together and cross checks their findings.
C.4		Review documentation	The DOCUMENTATION that the REPORTING ORGANISATION has provided is reviewed. Any FINDINGS (positive or negative) are registered.
C.5		Check adherance to method	The adherance to the method that the REPORTING ORGANISATION has employed is assessed.
C.6		Check conformity company with legal requirements	The conformity of the company to legal requirements is assessed. Examples might be legal requirements such as minimum wage, permit requirements or work safety provisions.
C.7		Check continious improvement of environmental performance	It is checked whether the REPORTING ORGANISATION improves their environmental performance continuously over the years.
C.8		Communicate with organisation concerning findings	The AUDITOR communicate with the organisation to discuss any FINDINGS. The main goal is to clarify any doubts that the AUDITOR has.
C.9		Require organisation to take corrective action	The REPORTING ORGANISATION has to take CORRECTIVE ACTION on any material misstatements that the AUDITOR has found.
C.10		Require evidence of corrective action	The REPORTING ORGANISATION has to provide EVIDENCE of the CORRECTIVE ACTION. This might, for example, include evidence that the REPORTING ORGANISATION has edited their NON-FINANCIAL REPORT.
C.11		Evaluate evidence of corrective action	The AUDITOR assesses the EVIDENCE that the REPORTING ORGANISATION has provided on the CORRECTIVE ACTION.
C.12		Investigate whether other sites are affected	In the case that the REPORTING ORGANISATION has multiple SITES, and the AUDITOR has found FINDINGS on one of those SITES, he investigates whether other SITES are affected.
C.13		Conduct closing meeting	A closing meeting is conducted by the AUDITOR and the REPORTING ORGANISATION. The main goal is to discuss the final FINDINGS of the AUDITOR.
D.	Report results		
D.1		Create audit report	The AUDITOR creates the AUDIT REPORT, which includes all the FINDINGS.
D.2		Form assurance conclusion	The AUDITOR forms the ASSURANCE CONCLUSION, in which he decides whether he wants to provide ASSURANCE on the NON-FINANCIAL REPORT.
D.3		Provide assurance statement	The AUDITOR provides ASSURANCE on the NON-FINANCIAL report by means of an ASSURANCE STATEMENT.

Figure 26: Table of activities of Generic Model

ID	Description	Explanation	Type	Concept	Role	Owner	Priority
Epics and User Stories for Audit extension openESEA							
		Story details		Meta-model (V4.2) traceability			
US-I-0	Initiate audit		Epic				
US-I-1	As a network administrator, I want to select organisations that have to be audited as part of a campaign.	A campaign contains several organisations that create an ESEA account through a single method. Auditing can also be part of this campaign. In a campaign, the network administrator should be able to start a batch audit process and select the organisations that have to be audited.	User story	ACCOUNT AUDIT	NA	Ties	Must
US-I-2	As a network administrator, I want to receive recommendations on organisations that have to be audited.	A network administrator receives a list of organisations based on a automated process where the recommendation finds place earlier. Based on this shortlist the network administrator can start the audit process for the organisations. The recommendation can be done by performing US-I-2A-B-C	User story	ACCOUNT AUDIT	NA	Artur	Must
US-I-2A	As a network administrator, I want the recommendations to be based on outlier detection.	Outliers need to be identified in two ways: (1) comparing the organisation to peers, (2) comparing to previous accounts of same organisation. This leads to a presentation of different orgaisations based on the fact that their account (or specific information about the account) is detected as an outlier.	User story	ACCOUNT AUDIT	NA	Artur	Must
US-I-2B	As a network administrator, I want the recommendations to be based on borderline certification score.	If the certification has been achieved just on the edge, e.g. that a organisation achieved 5.1 points, where companies are certified if score > 5, a indicator with the most impact that made the company achieving this certification will be recommended to audit.	User story	ACCOUNT AUDIT	NA	Artur	Must
US-I-2C	As a network administrator, I want the recommendations to be based on scoring risk (details to be added).	This includes the calculation of the scoring system where the absolute values are calculated for each indicator, and based on a specific case where an organization is on the border of getting, (or not getting) a certification, the scoring system would make a recommendation on which indicator should be audited.	User story	ACCOUNT AUDIT	NA	Artur	Must
US-I-2D	As a network administrator, I want the outliers to be presented through a visualization method.	This visualization could contain a graph with the specific outlier, a boxplot per indicator, a histogram of the indicators per account or a scatterplot	User story	ACCOUNT AUDIT	NA	Artur	Could
US-I-3	As an organisation administrator, I want to be informed when my organisation is selected for audit via email.	When the audit is triggered or started by the NA in US-I-1, the organisation that is going to be audited will need to be informed.	User story	ACCOUNT AUDIT	OA	Ties	Could
US-I-4	As an organisation administrator, I want to start an audit process.	An organisation adminstrator should be able to start an audit process for his company. This is an individual audit process for a single ESEA account.	User story	ACCOUNT AUDIT	OA	Ties	Must
	I want to be able to have multiple audit processes		User story			Ties	Could
US-P-0	Prepare audit		Epic				
US-P-1	As a network administrator, I want to assign an auditor.	When a batch audit process has been started (US-I-1), the network administrator should be able to assign an auditor for each individual audit process that is part of the batch.	User story	ACCOUNT AUDIT	NA	Ties	Must
US-P-2	As an organisation administrator, I want to assign an auditor.	When an individual audit process has been started (US-I-4), the organisation administrator should be able to assign an auditor for the process.	User story	ACCOUNT AUDIT	OA	Ties	Must
US-P-3	As an organisation administrator, I want to set an end date for the audit process.	The organisation administrator should be able to set an end date for the process. This end date can be changed at a later moment by the organisation administrator, and is merely a suggestion and not a restriction.	User story	ACCOUNT AUDIT	OA	Ties	Should
US-P-3A	As an organisation administrator, I want to set the assurance level for the audit process.	The goal of some audit processes is for the auditor to provide assurance on the audited information. If this is the case, the organisation administrator sets the desired assurance level: reasonable, limited or no assurance. This is the desired assurance level, and at the end of the audit process, the auditor decides whether he wants to provide assurance at all.	User story	ACCOUNT AUDIT	OA	Ties	Must
US-P-4	**SSS:** As an auditor, I want to select questions to audit.	For a single stakeholder survey, the auditor wants to select which questions to audit. This might be based on recommendations (US-P-4).	User story	QUESTION RESPONSE	AU	Ties	Must
		Based on the recommendations that the auditor will receive from US-P-5A-B-C, the risks of a		QUESTION			

Table 7: Requirements part 1

US-P-4	**SSS:** As an auditor, I want to select questions to audit.	For a single stakeholder survey, the auditor wants to select which questions to audit. This might be based on recommendations (US-P-4).	User story	QUESTION RESPONSE	AU	Ties	Must
US-P-5	**SSS:** As an auditor, I want to receive recommendations on questions to audit.	Based on the recommendations that the auditor will receive from US-P-5A-B-C, the risks of a materiality misstatement, can be presented to the auditor.	User story	QUESTION RESPONSE	AU	Artur	Must
US-P-5A	**SSS:** As an auditor, I want the recommendations to be based on outlier detection.	Through an outlier detection algorithm, an outlier can be detected ad therefore recommended to the auditor. These outlier detection could be performed by different unsupervised machine learning algorithms.	User story	QUESTION RESPONSE	AU	Artur	Must
US-P-5B	**SSS**: As an auditor, I want the recommendations to be based on scoring risk.	This includes the calculation of the scoring system where the absolute values are calculated for each indicator, and based on a specific case where an organization is on the border of getting, (or not getting) a certification, the scoring system would make a recommendation on which indicator should be audited.	User story	QUESTION RESPONSE	AU	Artur	Must
US-P-5C	**SSS**: As an auditor, I want the outliers to be presented through a visualisation method.	This visualization could contain a graph with the specific outlier, a boxplot per indicator, a histogram of the indicators per account or a scatterplot. This presetation takes the single stakeholder survey responses into account (one way of visualising).	User story	QUESTION RESPONSE	AU	Artur	Must
US-P-6	**MSS**: As an auditor, I want to select the number of survey responses to audit out of a multi respondent stakeholder survey.	For a multi stakeholder survey, the auditor wants to select the number of survey responses to audit. Absolute number or percentage based.	User story	SURVEY RESPONSE	AU	Ties	Must
US-P-6A	**MSS**: As an auditor, I want to have an overview of the survey responses that have been selected: names and emails.	To be presented to the reporting organisation when doing this audit in person.	User story	SURVEY RESPONSE	AU	Ties	Must
US-P-7	**MSS**: As an auditor, I want to receive recommendations on survey responses to audit. Randomly selected (number or a percentage).	A auditor receives a list of survey responses based on the recommendations created in US-P-7A-C-C. Based on this shortlist this recommendations, the auditor can start the audit process for the organisations.	User story	SURVEY RESPONSE	AU	Artur	Won't
US-P-7A	**MSS:** As an auditor I want the recommendations to be based on outlier detection	Through an outlier detection algorithm, an outlier in the multi stakeholder survey can be detected and therefore recommended to the auditor. These outlier detection could be performed by different unsupervised machine learning algorithms.	User story	SURVEY RESPONSE	AU	Artur	Won't
US-P-7B	**MSS**: As an auditor, I want the recommendations to be based on scoring risk.	This includes the calculation of the scoring system where the absolute values are calculated for each indicator, and based on a specific case where an organization is on the border of getting, (or not getting) a certification, the scoring system would make a recommendation on which indicator should be audited.	User story	SURVEY RESPONSE	AU	Artur	Won't
US-P-7C	**MSS**: As an auditor, I want the outliers to be presented through a visualisation method.	This visualization could contain a graph with the specific outlier, a boxplot per indicator, a histogram of the indicators per account or a scatterplot.	User story	SURVEY RESPONSE	AU	Artur	Won't
US-CS-0	Conduct audit (single stakeholder survey)		Epic				
US-CS-1	As an auditor, I want to request documentation regarding question responses	To audit the question responses, the auditor needs documentation that supports the question response. The auditor wants to request this documentation from the reporting organisation. The auditor should be able to attach a message to the request.	User story	DOCUMENTATION	AU	Ties	Must
US-CS-2	As an organisation administrator, I want to upload documentation regarding question responses	After a documentation request (US-CS-1), the organisation administrator responds to the request by uploading documentation. The OA should be able to attach a message to the documentation. Should be left open until audit is finished.	User story	DOCUMENTATION	OA	Ties	Must
US-CS-3	As an auditor, I want to view the documentation on question responses	After documentation is uploaded, the auditor wants to view this documentation to verify the question response.	User story	DOCUMENTATION	AU	Ties	Must
US-CS-4	As an auditor, I want to request a correction of the question response when I find a misstatement	When the auditor identifies a misstatement, he might want to request a correction of the question response.	User story	CORRECTION REQUEST	AU	Ties	Must
US-CS-5	As an organisation administrator, I want to receive a notification via email that I have to correct a question response	When a correction has been requested, the OA wants to receive a notification that he has to correct a question response.	User story	CORRECTION REQUEST	OA	Ties	Should
US-CS-6	As an organisation administrator, I want to correct a question response when this is requested	When a correction has been requested, the OA wants to be able to correct the response of the question	User story	CORRECTION REQUEST	OA	Ties	Must

Table 8: Requirements part 2

US-CS-1	As an auditor, I want to request documentation regarding question responses	To audit the question responses, the auditor needs documentation that supports the question response. The auditor wants to request this documentation from the reporting organisation. The auditor should be able to attach a message to the request.	User story	DOCUMENTATION	AU	Ties	Must
US-CS-2	As an organisation administrator, I want to upload documentation regarding question responses	After a documentation request (US-CS-1), the organisation administrator responds to the request by uploading documentation. The OA should be able to attach a message to the documentation. Should be left open until audit is finished.	User story	DOCUMENTATION	OA	Ties	Must
US-CS-3	As an auditor, I want to view the documentation on question responses	After documentation is uploaded, the auditor wants to view this documentation to verify the question response.	User story	DOCUMENTATION	AU	Ties	Must
US-CS-4	As an auditor, I want to request a correction of the question response when I find a misstatement	When the auditor identifies a misstatement, he might want to request a correction of the question response.	User story	CORRECTION REQUEST	AU	Ties	Must
US-CS-5	As an organisation administrator, I want to receive a notification via email that I have to correct a question response	When a correction has been requested, the OA wants to receive a notification that he has to correct a question response.	User story	CORRECTION REQUEST	OA	Ties	Should
US-CS-6	As an organisation administrator, I want to correct a question response when this is requested	When a correction has been requested, the OA wants to be able to correct the response of the question.	User story	CORRECTION REQUEST	OA	Ties	Must
US-CS-7	As an auditor, I want to view the correction of the question response	The auditor wants to be able to view the corrected question response, including the old and the new value.	User story	CORRECTION REQUEST	AU	Ties	Must
US-CS-9	As an auditor, I want to be able to request additional documentation	An auditor might decide that the provided documentation is not sufficient to verify the question response. In this case, he might want to request additional documentation. This triggers US-CS-2 again. ALTERNATIVE: leave US-CS-2 open until audit is finished. This way, additional documentation can always be uploaded.	User story	QUESTION RESPONSE AUDIT	AU	Ties	Won't
US-CS-8	As an auditor, I want to be able to reject a question response	When the auditor identifies a misstatement, and the organisation is not willing to correct it, the auditor might want to reject this question altogether.	User story	QUESTION RESPONSE AUDIT	AU	Ties	Must
US-CS-9	As an auditor, I want to verify the question response	The auditor wants to verify the question response, in case he deems the documentation sufficient to support the question response.	User story	QUESTION RESPONSE AUDIT	AU	Ties	Must
US-CM-0	Conduct audit (multi stakeholder survey)		Epic				
US-CM-1	As an auditor, I want to invite respondents to verify their responses	The respondents of the survey responses selected in US-P-6 will be invited to verify their respones in person.	User story	SURVEY RESPONSE AUDIT	AU	Ties	Must
US-CM-2	As a survey respondent, I want to view my survey response	After being invited (US-CM-1), the respondents will verify whether the survey responses are indeed theirs. They will receive a link on their email. When clicked, they will be presented with their survey response.	User story	SURVEY RESPONSE AUDIT	SR	Ties	Must
US-CM-3	As an auditor, I want to verify the selected survey responses	Three buttons: send email, verify, reject	User story	SURVEY RESPONSE AUDIT	AU	Ties	Must
US-CM-4	As a survey respondent, I want to verify my survey response	When viewing their survey response (US-CM-2), the survey respondents wants to verify that the question responses in the survey response have been entered by themselves. At the bottom of their survey response there is an option to verify this.	User story	SURVEY RESPONSE AUDIT	SR	Ties	Could
US-R-0	Report results		Epic		AU	AU	
US-R-1	**SSS:** As an auditor, I want the audit results to be visualised.	When all question responses/survey responses have been audited, the auditor wants to finish the auditing and have the results be visualised. This contains all audited questions and their results, and, where applicable, corrections.	User story	ACCOUNT AUDIT	AU	Ties	Must
US-R-1A	**MSS:** As an auditor, I want the result of a multi stakeholder survey audit to be visualised	Visualised as in: verified or not verified	User story	SURVEY RESPONSE AUDIT	SR	Ties	Should
US-R-2	**SSS:** As an auditor, I want to form an assurance conclusion.	After all question responses/survey responses have been audited, the auditor mights want to form an assurance conclusion. In this case, the auditor decides whether he wants to provide assurance or not.	User story	ACCOUNT AUDIT	AU	Ties	Must
US-R-3	**SSS:** As an auditor, I want to create an assurance statement.	When the auditor has elected to provide assurance, he wants to create an assurance statement.	User story	ACCOUNT AUDIT	AU	Ties	Must

Table 9: Requirements part 3

D. Screenshots of OpenESEA tool

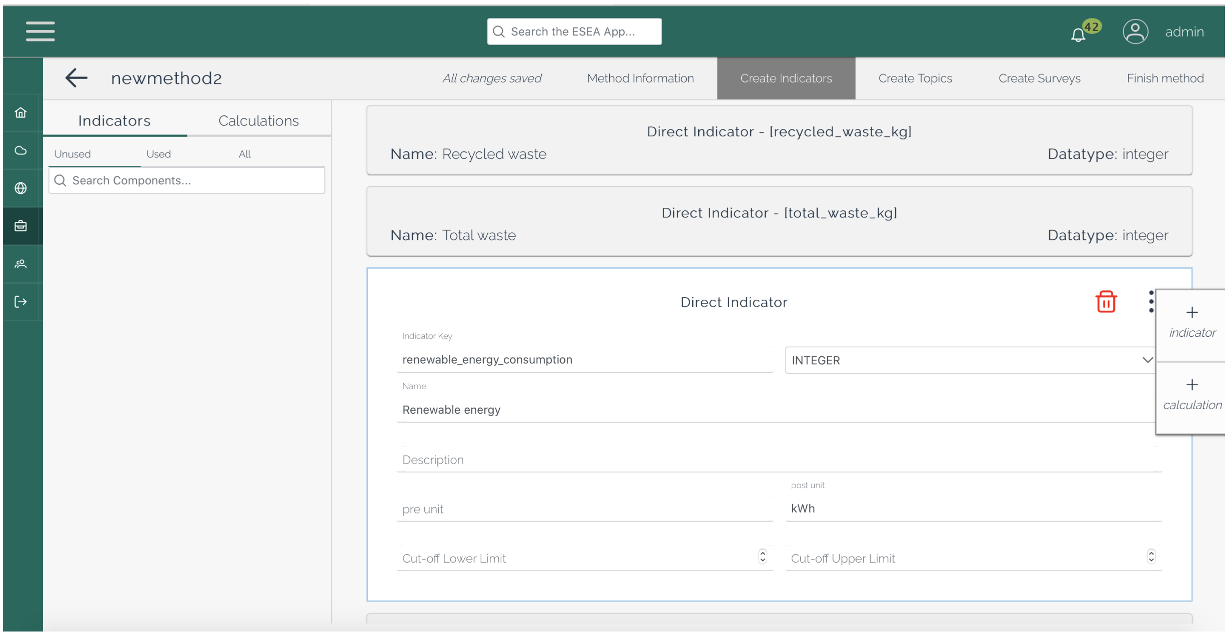

Figure 27: OpenESEA Cut-Off Lower and Upper limit

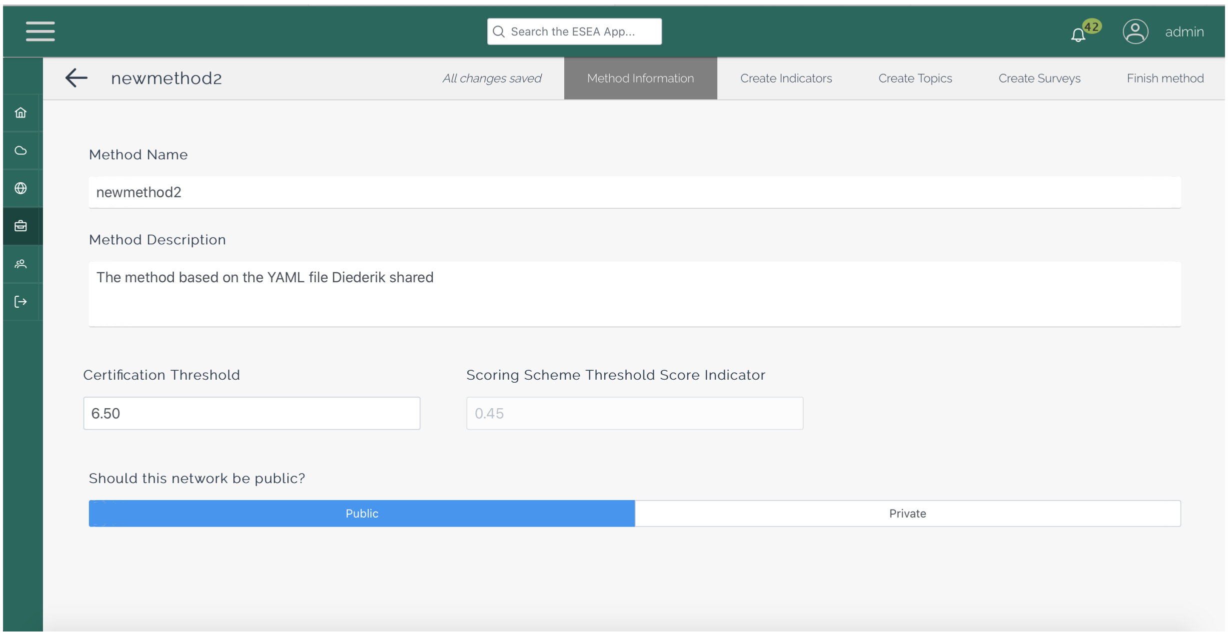

Figure 28: OpenESEA Scoring scheme: method specification

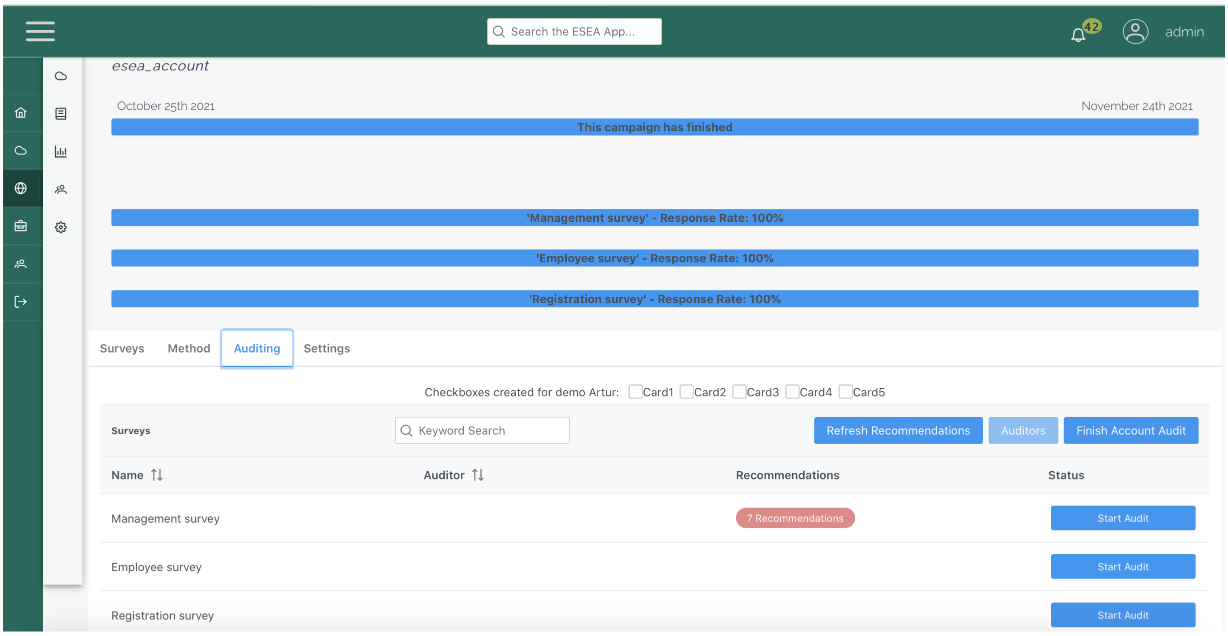

Figure 29: OpenESEA Scoring scheme: Recommendations

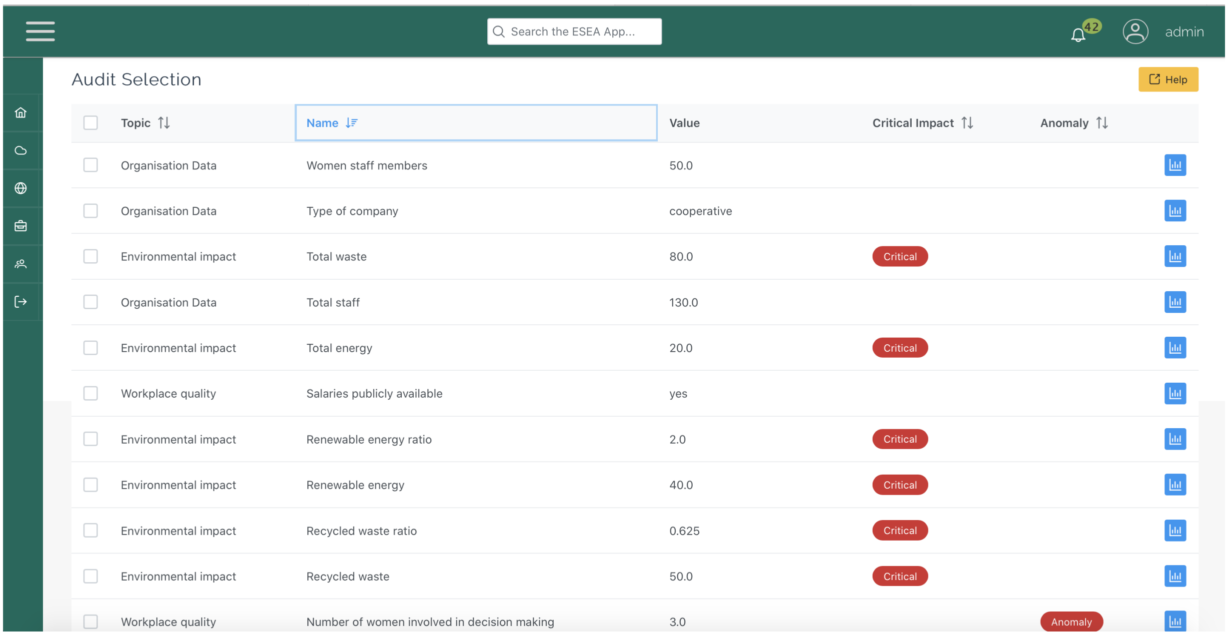

Figure 30: OpenESEA Scoring scheme: list of indicators for audit

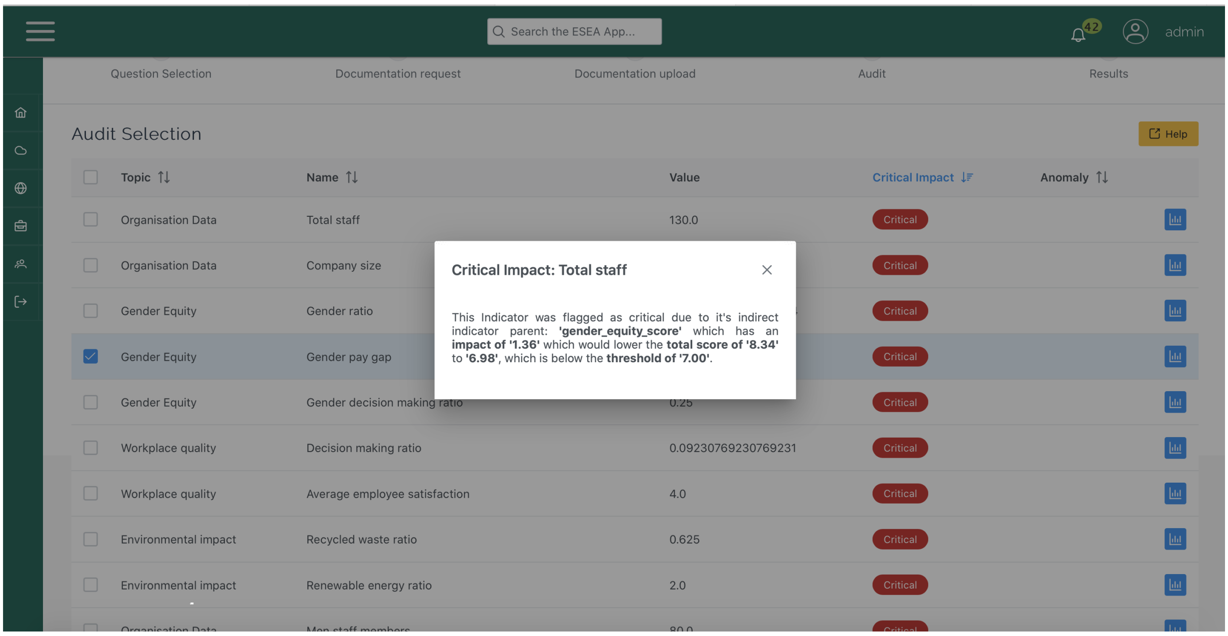

Figure 31: OpenESEA Scoring scheme: Critical value and its parent indicator