Navigating the Sustainability Reporting landscape – Method Specification and tool for Materiality Assessment

Rim Mukhopadhyay

Department of Information and Computing Sciences,
Utrecht University

Abstract. This work delves into the pain points encountered in the current industry practices for conducting materiality assessment, and aims to improve the process by reducing effort and subjectivity involved therein. The work consists of a method specification based on a case study. A part of the envisioned solution is implemented by creating a tool to automate parts of the process. Apart from reducing effort and subjectivity, the proposed solution incorporates weightage-schemes in creation of term document matrix, where the previous process considered only occurrences. The tool is validated to favorable outcomes from practitioners and researchers, with respect to being a valid solution and potential commercial implementation. The automation achieves to reduce the effort to less than 2% of the original. Future work in this line of research is improving the precision and recall parameters, social media sentiment analysis, and a learning module to further streamline and perfect the tool-output.

1   Introduction

Sustainability Reporting is a means by which organizations account for their socio-environmental impact on their surroundings, operations, and society as a whole. This not only ensures a sustainable brand image, but also engages employees, investors, and other stakeholders by engaging them in the operations and impact associated with the entire value-chain. Investor demand for more disclosure, transparency, accountability, global reporting framework synergies, and financial market convergence are some of the driving factors behind organizations adapting to Sustainability Reporting. Gray (2000), aptly traces this need saying that the complex, civilized society of today or entities therein, aiming for a sustainable future with less-exploitative relationships with the environment, undoubtedly needs to steer in the direction of socio-environmental reporting.

Sustainability reporting helps organizations set goals, measure performance, and manage change to make their operations more sustainable. A sustainability report conveys disclosures on an organization's impacts – be they positive or negative – on

the environment, society, and the economy. In doing so, sustainability reporting makes abstract issues tangible and concrete, thereby assisting in understanding and managing the effects of sustainability developments on the organization's activities and strategy. This gives rise to the need for standard, uniform norms to measure, report and communicate organizations' inclination to corporate conscience.

Internationally agreed disclosures and metrics enable information contained within sustainability reports to be made accessible and comparable, providing stakeholders with enhanced information to inform their decisions. The paper by Manetti and Becatti, (2009), points out the reliability and comparability issues that exist in the current arena of Sustainability Reporting. Gray (2010) too affirms the existence of this gap existing in the arena of sustainability reporting, which makes it difficult for such reports to be used by the target stakeholder groups.

The Global Reporting Initiative promotes standards and is being widely adopted in the field of sustainability reporting. According to The Global Reporting Initiative, "defining what matters" is the first step to increase the visibility of the sustainability reports. This would enable comparison between reports, which would contribute to the visibility factor. The process of identifying and determining which topics should be prioritized for reporting, is done by "Materiality Assessment".

However, the range of topics that might be reported on is vast, with blurry edges. Additionally, the interests of respective stakeholder groups (investors, customers, regulatory bodies) differ by sector. What is important to an automobile manufacturing organization, is very different from that of a food and beverage maker. Currently, practitioners undertake Materiality Assessment exercises manually, as observed in our case study. This leads to immense effort consumption, redundant work-cycles, and a large amount of subjectivity, owing to variable interpretations of the document set used to identify material topics.

It is thus required to make such reports more uniform, standard, and comparable. Our work proposes a method and a tool to automate the Materiality Assessment process, thereby eliminating the subjective aspects to a large extent, and minimizing the most time-consuming and redundant parts of the process.

By usage of this tool, we aim to resolve the issue "*To date, few studies have inquired into quantitative methods to support materiality assessment in sustainability reporting, and these have not addressed the issues of subjectivity or of completeness in the reporting.*" Costa et al (2016).

This tool and the framework behind it would benefit Academia by making clear distinctions regarding the exact definitions, simplify the objective, methodology, impact, end user, etc. of the concepts involved. Commercially, it will aid businesses, who are required to include non-financial statements in their annual reports from 2018 onwards, as per Directive EU Law (2014/95/EU). Disclosures are required for environmental protection, social responsibility and treatment of employees, respect for human rights, anti-corruption and bribery, diversity on company boards, in terms of age, gender, educational and professional background. This tool streamlines the current process followed in the industry for Materiality Assessment. It would also help businesses to stay ahead of reporting trends in their respective industries, and chart their strategic courses.

This paper starts with stating the methods and techniques that have been used in this research and the research life-cycle. Next we investigate the nuances of materiality assessment practices by means of a case study. The findings of this investigation help us formalize the problem scenario by creating method specifications, and define the requisites for the solution. Then we proceed to the tool design and implementation specifics. This section also showcases some of the outputs of the tool. Next to this, we validate the effectiveness and efficiency of our solution in resolving the problem scenario. The validation results are discussed; it depicts the solution as perceived favorably by the stakeholders, and the time-savings. Some results could be improved by future enhancements to the tool, and these are discussed in the concluding section. The tail of the paper consists of the appendices containing the evidences and technicalities of the preceding sections.

2 Research Method

This section delves into the methods and techniques applied during this research. We discuss the modeling techniques and notations that are encountered throughout this paper. This is followed by a discussion of the design-cycle and lifecycle, and explanation of the steps therein. The deliverables and metrics have been briefly discussed as well.

2.1 Diagrammatic Representation

To represent the processes involved in this project, we shall make use of the Process Deliverable Diagram (PDD) technique, as proposed by van de Weerd and Brinkkemper (2008). A PDD shows the activities on the left hand side, and the corresponding concepts and deliverables on the right hand side. For our research, we have incorporated an additional notation to highlight the pain-points in the process. Each PDD is accompanied by an activities table and a concept definition table, which are found in the appendix. Other forms of illustrations are accompanied by respective legends.

2.2   Design Cycle

This research will be undertaken by adopting the iterative problem-solving method proposed by Wieringa (2009). Upon the design and development of the deemed apt solution, the next phase "Treatment Validation" is started, where the solution is tested against the initial problem scenario. The fourth phase of Wieringa's cycle, the Treatment Implementation, is not carried out as this is a short-term project. Figure 1 below is the overall set of Activities and their corresponding outputs/deliverables for the application of Wieringa's Design Cycle to our Project, for a high-level overview.
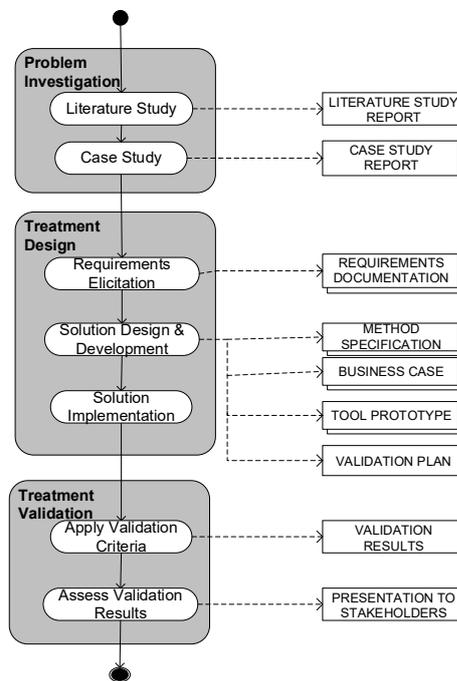

Fig. 1. Overview of research methodology

The sequence of steps followed is elaborated below. The first level headings (1,2, 3..) denote the Design Cycle steps, and the subheadings under these are the steps and techniques followed therein.

1.   Problem Investigation
a) Literature Review – A snowball study of the academic documentation, to further verify the existence of the problem, and uncover existing research. Google Scholar and other academic databases like DBLP have been used for this. The summarization of the learnings can be seen in Appendix 1.

b) Case Study –   A sustainability-strategy-consulting company "Sustainalize" was contacted in Utrecht, the Netherlands, who perform Materiality Assessment for their customers in various sectors. The case study at Sustainalize comprised of:
   i. Field Observation – Observation of consultants in Sustainalize conduct real-time Materiality Assessment for their Customers.
   ii. Apprenticeship – Conducting Materiality Assessment exercise under the supervision of consultants at Sustainalize, with the aim of uncovering problem-areas/ pain-points.
   iii. Data Collection – Semi structured, open interviews with consultants of Sustainalize to get deeper understanding of the nuances of the process. Interview notes may be found in the Appendix 3. Access to the documents relevant for conducting Materiality Assessments was obtained (residing in the company-repository (Google Drive)).

2. **Treatment Design**
   The treatment comprised of the following deliverables:
a) Method Specification – Capturing the Materiality Assessment method and formal documentation, in the form of Process Deliverable Diagrams and other models.
b) Solution/ Tool Design, Development & Implementation - In this phase, we have applied Agile Development methodology, to produce multiple increments of the solution to arrive at a final working version.
   i. Technical collaboration – The development was carried out in collaboration with a student from Utrecht University. The tool was developed using R Studio, making use of the predefined text-mining packages available therein.
c) Business Case – The feasibility study of the commercial viability was explored using SWOT Analysis. The Business Models were discussed with the project-owners at Sustainalize. Owing to the short tenure and academic nature of this research, it was agreed that future enhancements (after the initial working version) and commercialization would be taken up by Sustainalize and their development partners, Eindhoven based F19.

3. **Treatment Validation**
   For the validation, "Actual" and "Perceived" parameters were captured.
a) For the Actual parameters, the measures were:
   i. Actual Efficiency – Measured by comparing time consumption of the current processes (manual materiality assessment) versus the time consumption after applying the treatment (using the tool).
   ii. Actual Effectiveness – Measured by Precision and Recall metrics, as proposed by Ting (2011). The metrics tailored to the situation are:

   Precision = Total number of items (material topics) retrieved that are relevant/Total number of items (material topics) that are retrieved.

   Recall = Total number of items (material topics) retrieved that are relevant/Total number of relevant items (material topics) in the input sources

b) For the perceived parameters' validation, the "Method Evaluation Model" proposed by Moody (2003) was used, tailored to the situation at hand. The modified model is shown in figure 2.
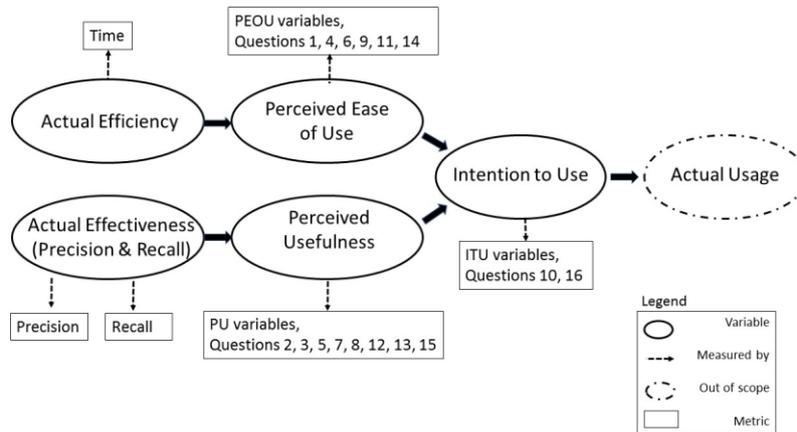
Fig. 2. Method Evaluation Model for the solution

Independent variable is the treatment designed (tool) for the problem. Three dependent variables were used to measure the outcome.
The dependent variables measured were:
 i.   Perceived Ease of Use (PEOU) – Measured by 6 survey questions
 ii.   Perceived Usefulness (PU) – Measured by 8 survey questions
 iii.   Intention to Use (ITU) – Measured by 2 survey questions.
The Fourth variable (Actual Usage) proposed by the original model was deemed out of scope owing to time constraints of our project, and keeping in mind the design cycle followed herein. The questionnaire pertaining to this model was adapted from Davis et al's (1989) study, with changes tailored according to our subject, and circulated in an online Survey format. Responses were obtained in a five-point Likert scale. Details of the tailored questionnaire, and responses are available in the Appendix 5.

.

3   Problem Investigation

The Problem Investigation was conducted by means of Case Study and Situation Study at a sampled Business. To gain insight into the various aspects of the problems, the following models were created.

1.   Context Model - For the basic problem and context understanding, the DPSIR Model was used. It uses five constructs (Driving Forces, Pressure, State, Impact, and Response) to capture interactions of environmental pressures on human and societal entities. The reason we have selected this model is that has been adopted by the European Environmental Agency (EEA) and the European Statistical Office in 1997. This model was then superposed on the problem at hand, to understand each of the nodes. The situational attributes for each of the nodes were then identified, which led to the Situational DPSIR Model, shown in figure 3 below.
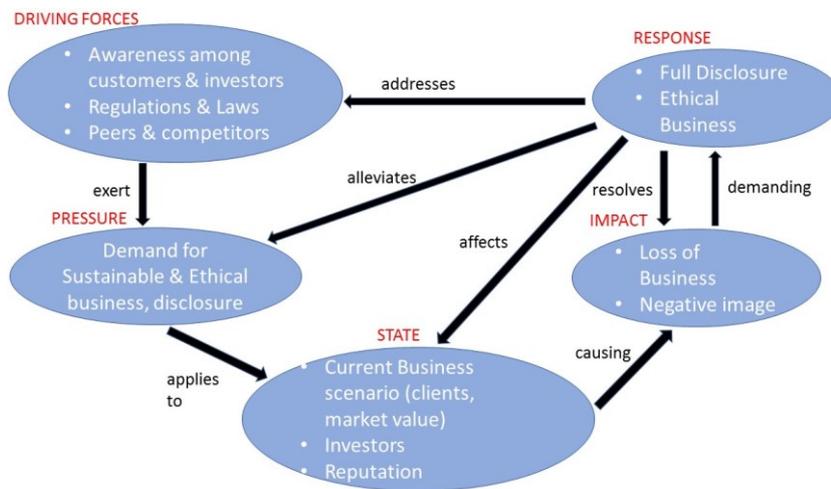

Fig. 3 – DPSIR Model – situational to Sustainability reporting

The mapping of the generic and situational DPSIR models is tabulated in the Appendix 2. The effect of each node on other nodes is depicted by means of relationships (on the arrow-connectors). The starting point is the Driving Forces, that cause pressure, threatening to damage the business, and are required to be acted upon.

2.   Stakeholder Model – To understand who are impacted by the phenomenon of Sustainability Reporting (specifically Materiality assessment), we have created the following Model. This model depicts how each of these stakeholders would benefit from this model.

Fig. 4 – Stakeholder Model

Each of these stakeholder categories interact with Sustainability reports with varying perspectives, as shown above. These stakeholder groups were arrived at based on review of the documents that serve as input to the Materiality Assessment process. Evidence of this traceability may be found in the Appendix 2.

3.   Conceptual Model – The prevalent concepts in this research have been explained in the model below. It shows the prevalent actors, concepts and the process involved in the Materiality Assessment paradigm. The model also attempts to explain the sequence of triggers and result of the processes discussed herein.
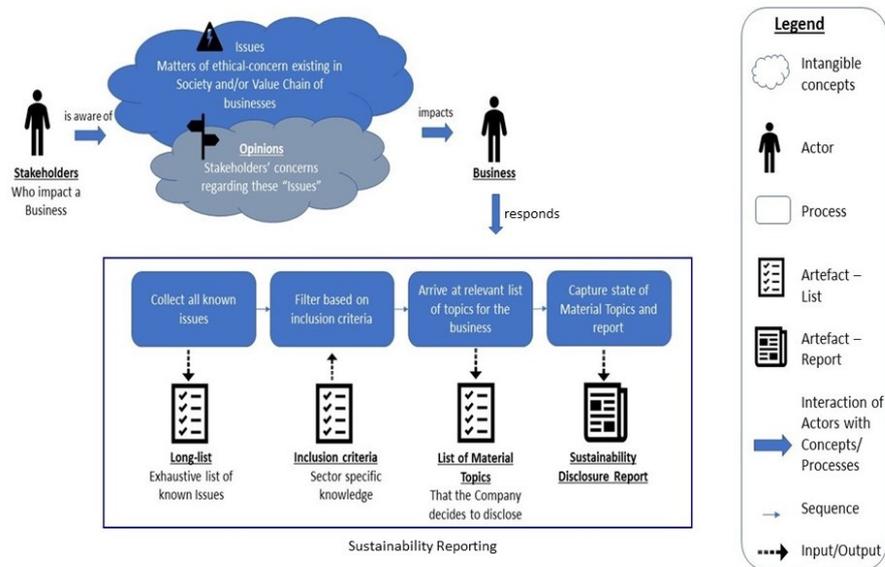
Fig. 5 – Conceptual Model

A stakeholder belongs to any of the groups shown in the Stakeholder Model in fig. 4. The stakeholders' opinion about socio-environmental issues makes a business susceptible to threats, shown in fig.3, DPSIR Model. This prompts businesses to take certain action to alleviate the pressure and appease the stakeholder concerns. This response, in our case, is the Sustainability Disclosure, of Materiality Assessment is the starting point.

4. Process & Deliverable Model - To understand the current process of Materiality Assessment, a step by step Process Deliverable Model was created. The current process followed by Sustainalize is represented via the following PDDs. Each activity has a unique identifier. Figure 6 shows us a high-level view of the overall Materiality Assessment Process to facilitate a basic understanding. We have skipped the deliverables in this figure on purpose, as this serves to explain the activities only. The subsequent PDD delves into each of the steps, details the sub-activities and deliverables, and explains the relations/ dependencies between the deliverables. It also highlights the pain points in the current process.
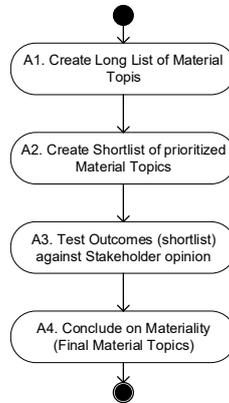
Fig. 6 – High Level view of Materiality Assessment Process

We have identified two types of pain points, as distinguished by the color.

Primary Pain Points (denoted in Red) are the ones we propose to resolve initially, as these cost the company the most effort. Also, the subjectivity involved in document analysis and interpretation tasks could be removed. These steps are carried out manually for every customer, and there is a huge degree of redundancy involved. Resolving these pain points would alleviate the problem scenario to a large extent (measures are available in the Validation Criteria section).

Secondary Pain Points (denoted in Orange) translate into the "nice to have features" in the requirements (and subsequently functional specifications). These should be interpreted as "improvement opportunities" for the current process. We try to resolve these based on time and resource availability. This prioritization has been done based on the amount of time spent on the activities. For Example, the first Primary Pain Point, "Conduct Desk Research" is estimated up to 30 hours of manual effort. It is to be noted that the pain-points have been arrived at based on the estimates provided to the customers by Sustainalize. For confidentiality reasons, these estimates are not disclosed in this document, as they contain confidential business-information. The pain-points relating to matters of subjectivity and accuracy have been identified based on interviews. Also, the researcher has undertaken the Materiality Assessment process for a few customers to get a first-hand experience of which tasks face these issues of subjectivity/accuracy.
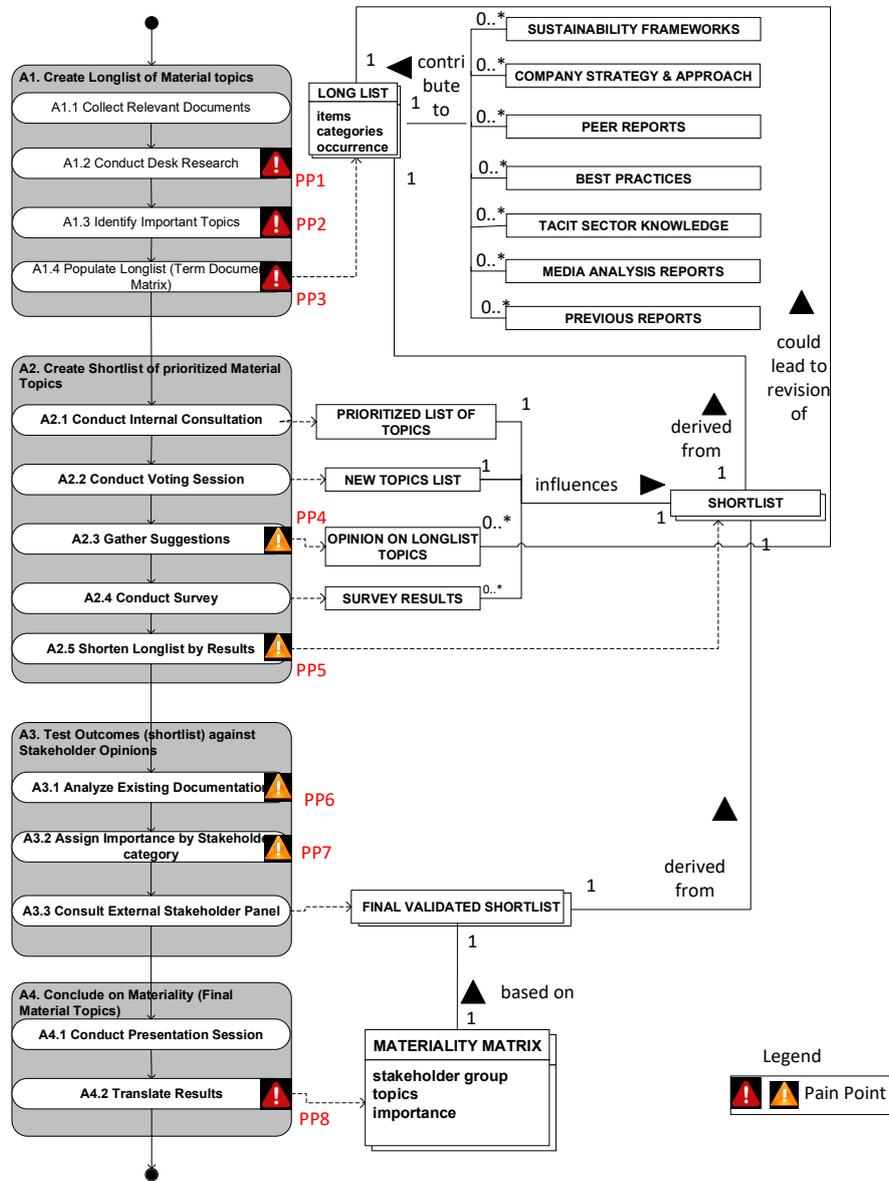
Fig. 7 – Current Materiality Assessment Process, detailed

The steps in the above PDD are explained below. Elaboration of the activities and deliverables in the above PDD may be found in Appendix 2.

A1: Create Longlist of Material Topics - The goal of this step is to arrive at a populated Term Document Matrix, which would serve as the input to create a shorter list of relevant topics (based on prioritization), which in a later stage can be presented to both internal (step 2) and external stakeholders (step 3) for further assessment and ranking. A sample Longlist is shown below.

Categories	Long list of matters	Score	GRI 4.0	GRI Sector Industry food Processing	SASB (general)	SASB (industry specific)	ISO 26000	Wessanen Annual Report	Hain celestial ANN...
	Bio-based economy	1	X						
	Investments and research	1							
	Application of biotechnology	1	X						
Business	Educating customers/consumers	2						X	X
	Employee health & safety	11	X	X	X	X	X	X	X
	Occupational health and safety risks	6	X		X				
	Labor management systems	0							
	Employee well-being/ wellness	5			X	X		X	X
	Work-life balance	3		X					
	Labor relations and union practices/ social dialogue	6	X		X	X	X		
	Freedom of association and collective bargaining	5	X			X			
	Conditions of work and social protection	4			X	X	X		
/ Labor relations	Strikes, lock-outs and/or labour unrest	1	X						
	Diversity and equal opportunity	6	X		X				X

Fig. 8 – Longlist

The arrow "1" shows the "Longlist of matters" column, where the exhaustive list of topics is listed. The arrow "2" shows the input documents. For each cell in the "Longlist of matters" column, the consultant tries to identify a match in each of the documents mentioned in the row of documents. The categories of these input documents vary from frameworks to peer reports to other documents from the value chain. The arrow "3" is the body of the Term Document matrix, wherein an "x" is marked for each matching term in the respective document.

Pain-points identified:
PP1 at A1.2 – Conduct Desk Research
This involves reading a multitude of documents and reports, and is a pain-point because of the time required.
PP2 at A1.3 – Identify important topics
Based on the documents read, the identification of material topics is done, and this is a pain-point because of the subjective nature of the task. It could yield different results depending on the person, their understanding of the topic/language, and productivity level at any given point in time.
PP3 at A1.4 – Populate Longlist (Term Document Matrix)
The task of manually populating the Longlist has been identified as a pain-point because of the time requirements and accuracy-margins. This is a tedious task, and interviews reveal that the person conducting this finds a lack of focus after getting halfway through the matrix.

A2: Create shortlist of prioritized Material Topics
To prioritize the topics from the short list, the process starts with an internal consultation. This is done in an interactive voting session in which the shortlist from step 1 is discussed in detail. The session is aimed at assigning importance to each of the topics included in the shortlist from the viewpoint of the company undertaking the materiality assessment.

Pain-points identified:
PP4 at A2.3 – Gather suggestions, PP5 at A2.5 – Shorten Longlist by results
These two tasks, though not too time consuming, have been identified as pain-points by looking at the nature of them. Incorporating the results into the Longlist to exclude certain terms, can easily be automated, thus leading to a more streamlined process overall.

A3: Test outcomes (shortlist) against external stakeholder opinions
To complete the process of materiality assessment – and comply to the G4-guidelines – we need to test the outcomes of the preceding steps again against the opinions of external stakeholders. This is done based on an analysis of existing documentation.

Pain-points identified:
PP6 at A3.1 - Analyze existing documentation
This is identified as a pain-point because of the tedious and time-consuming nature of the task. There is also subjective-interpretation involved.
PP7 at A3.2 - Assign importance by stakeholder category
This step is a pain-point (or more of an improvement area), because this could be easily automated, and thereby cut down the time requirements.

A4: Conclude on materiality (final Material Topics)
The final step in their support is to combine the outcomes from all previous steps and translate this into a final overview of material topics.
Pain-points identified:
PP8 at A4.2 – Translate results
Reporting the results in a graphical format is something that could be easily taken over by a program, thereby reducing the time consumption. This is a redundant task for every customer, and there is not much human-expertise or opinion required here.
Each of these steps are elaborated in detail in Appendix 3.
The output of the whole process is summarized by the Materiality Matrix, shown below:

Fig. 9 – Materiality Matrix

The horizontal axis shows the importance stakeholders assign to certain topics; the vertical axis the importance COMPANY X International assigns to the topics. Topics in the upper right corner are perceived most significant and are expected to be addressed in the sustainability report.

In this chapter, we investigated the problem in depth, and identified the pain points or inefficient parts of the process of Materiality Assessment at Sustainalize. Next chapter deals with the resolutions to each of the problematic phenomena described above.

4   Treatment Design

The treatment proposed for the tool consists of an improved method, specifications for a tool, and implementation of the tool. This section explains the tool and the underlying methods.

Based on the problem investigation and the models constructed, the research now proceeds to creation of prototype, with an attempt to mitigate the pain points elaborated in the previous section. To set the expectations, a black-box model of the tool was created, and is shown in the figure below. It depicts the inputs and output of the tool. It is to be noted that there are multiple intermediary outputs (short-list, word-cloud) which have not been shown here for the sake of simplicity.
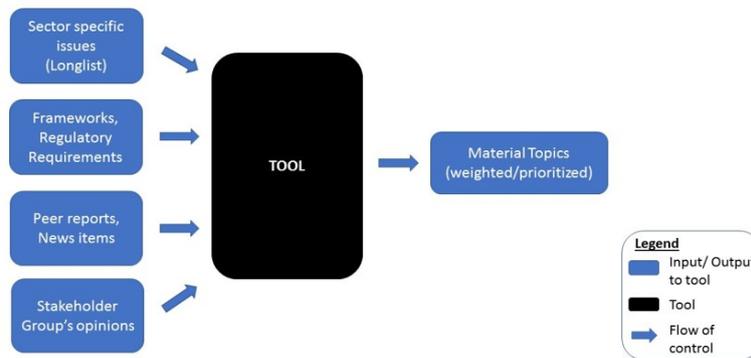
Fig. 10 – Black-box model of tool

The detailed requirements specifications, prioritizations, and decision-making rationale may be found in Appendix 3. Next, the Features Model was created based on the technique suggested by Brunetti & Golob (2000). This takes a look at the functionalities of the tool, and has been designed based on resolving the pain-points identified.
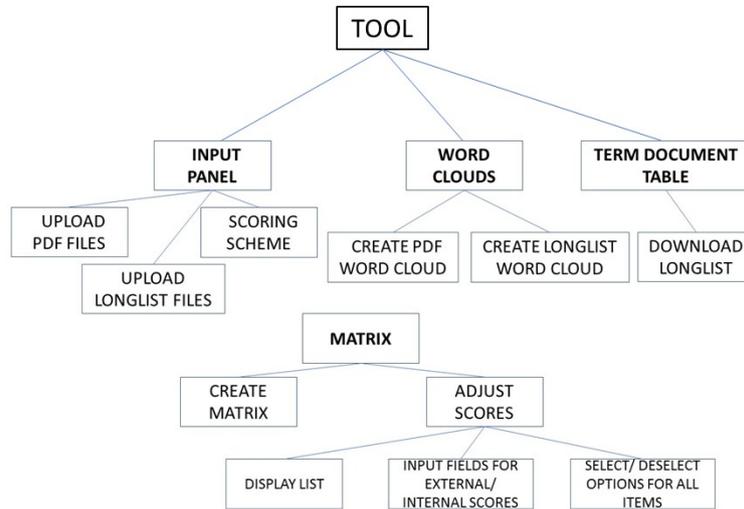
Fig. 11 – Features Model

We have added some additional features to the tool, which were not part of the current process, to provide the users with more options. These are discussed below.

• We have incorporated a feature to showcase the word occurrences in word-cloud format. This is meant to create a good impression for a commercial demonstration to customers and prospects. In PP2, the longlist is populated based on input document set (matches in longlist vs document). The tool enhances this by giving the user an extra option to find the most-occurring words in the documents without matching with the longlist. The aim of this step is to emphasize (to customers and prospects) the importance of the longlist in identifying material topics. These features are showcased in the screenshot below:

Fig. 12 – Screenshot of improvements, Word Cloud

• In PP3 – Populating the Term Document Matrix (Longlist). The original process captured the occurrences with an "x", wherever there was a match. The tool replaces this x with details of the occurrence, based on four schemes, mentioned below.

a) Count - Does topic occur in a PDF? Yes = 1, No = 0
b) Frequency - Number of occurrences of a topic in a PDF
c) Relative - (Frequency / Page count) > Threshold? Yes = 1, No = 0
d) Weighted - ((Frequency / Number of pdfs with frequency > 0) * total number of pdfs)

Fig. 13 – Screenshot of improvements, Scoring scheme

The figure below shows the tool output with details of occurrence.

Categories	Longlist	Score	Boskalis_AnnualReport_2016	CSR_Report_2016	DEME.Activity.Report 2016	DEME.fin_Jaarversl 2016
	Company competitiveness / competitive behavior / fair competition / anti-competitive behaviour	2	0	0	0	0
	Economic performance	1	0	0	0	0
	Long term viability of core business	0	0	0	0	0
	Market presence	0	0	0	0	0
	New markets / emerging economies	1	0	0	0,007	0,007
	Corporate governance / organizational governance	4	0,062	0,112	0	0,015
	Response to organizational change	0	0	0	0	0
	Board structure and independence	0	0	0	0	0
	Gender participation on governance bodies	0	0	0	0	0
	Executive compensation	0	0	0	0	0
	Business ethics	2	0	0	0	0
Business	Regulatory compliance	0	0	0	0	0
	Environmental compliance	0	0	0	0	0
	Regulatory and legal challenges	0	0	0	0	0
	Communications and engagement	0	0	0	0	0
	Shareholder engagement / Stakeholder engagement / Shareholder Expectation	4	0	0,038	0	0,037
	Innovation	6	0,062	0,05	0,169	0,022
	Innovation management process	0	0	0	0	0
	Sustainable innovations	0	0	0	0	0
	Circular economy	1	0,006	0,075	0	0
	Green entrepreneurial culture	0	0	0	0	0
	Bio-based economy	0	0	0	0	0

Fig. 14 – Tool output of Longlist

The improvements may be summarized as:
• New activity introduced - A1.35 – Select scoring scheme
• New attribute SCORE added to the concept LONG LIST
• New deliverable WORD CLOUD introduced

Further elucidation of the features may be visualized via screenshots in Appendix 4.
The traceability between the proposed features and pain-points is shown in the table below. With this, we show that all the pain points identified have been addressed by the tool.

Pain-Point number	Pain-Point name	Explanation	Requirement (to resolve)	Feature	Location (in tool)
PP1	Conduct Desk Research	Scanning through input documents, looking for links wih the Longlist	Document scanning to identify Phrase/term based on Longlist	Upload PDF files, Upload Longlist, Create PDF Word Cloud	Input panel, Tab - Word Clouds
PP2	Identify important topics			Text matching algorithm (runs in the backend)	Backend
PP3	Populate Longlist (Term Document Matrix)	For each item in Longlist, indicate occurrence matches with respect to each input document	Create Term Document Matrix based on previous step	Create Term Document Matrix	Tab - Term Document Table
PP4	Gather suggestions	Obtain opinion from Company about the topics (the matches found in previous step)	Option to input suggestions from company for all the Longlist matches	Create Longlist word cloud	Tab - Word Clouds
PP5	Shorten Longlist by results	Incorporate opinions from previous step in Longlist, thereby excluding some items and shortening the list	Filter out and shorten Longlist based on previous step	Option to select/ deselect Longlist items	Tab - Matrix
PP6	Analyse existing documentation	Scanning through second set of input documents	Document scanning to identify Phrase/term from shortened Longlist for second set of documents	Upload PDF files, Upload Shortlist	Input panel, Tab - Matrix
PP7	Assign importance by stakeholder category	For each topic, assign weightage based on previous step	Option to assign weightage to the shortlist of topics obtained from previous steps	Manual input to shortened list of topics	Tab - Matrix
PP8	Translate results	Plot Materiality Matrix	Select number of topics to display, and plot them on graph according to weightage	Create Matrix (plot graph)	Tab - Matrix

Table 1 - Traceability between Features and pain-points

To understand the afore-mentioned improvements, let us look at the following PDD. It is a modification of the previous detailed Materiality Assessment process PDD, with the improvements designed by our research. The improvements are denoted by the shaded boxes.
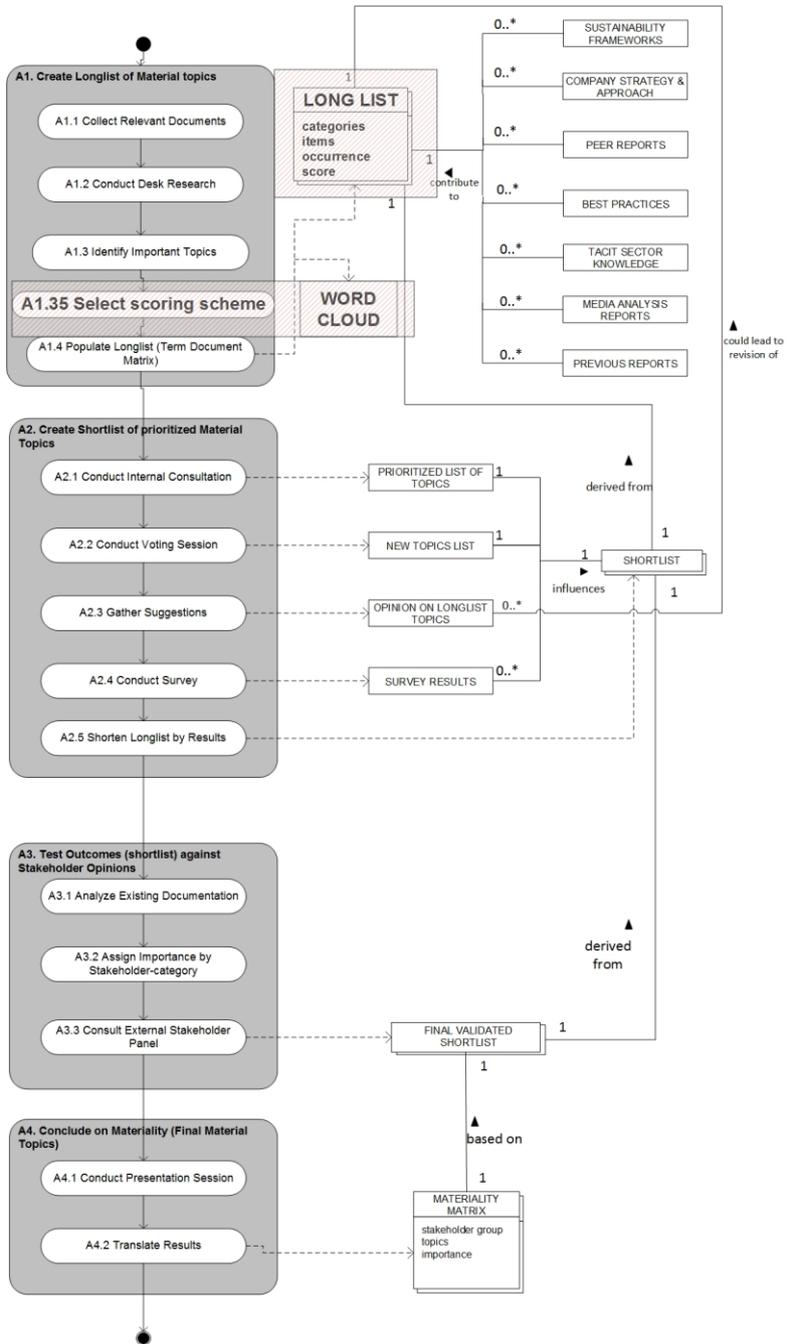
Fig. 15 – PDD with improvements

5 Treatment Validation

In this section, we look at the results based on the two aspects, the "Actual" and "Perceived" parameters.

5.1 Actual Parameters' Validation results
Two data sets were obtained, and the materiality Assessment exercise was performed on them manually, as well as with the tool. The results showed the following:
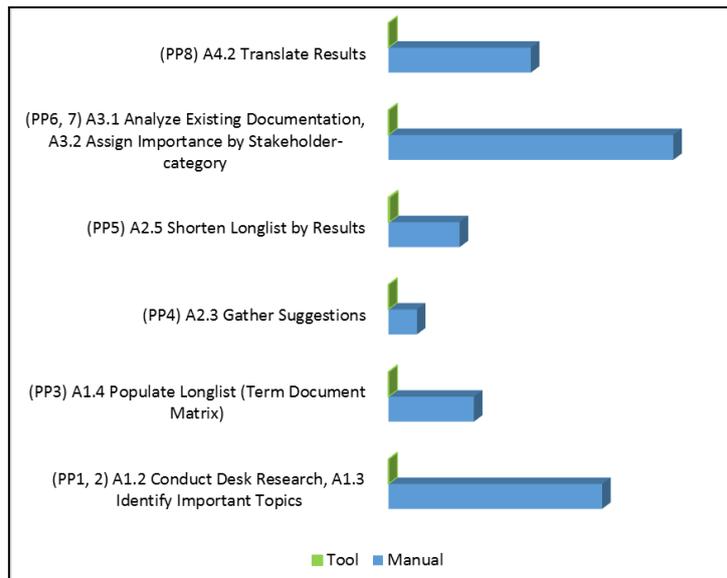
a. Actual Efficiency

Activity	Sub-activity	Data Set 1			Data Set 2		
		Manual	Using Tool	Difference	Manual	Using Tool	Difference
A1. Create Longlist of Material topics	A1.1 Collect Relevant Documents	NA	NA	NA	NA	NA	NA
	A1.2 Conduct Desk Research, A1.3 Identify Important Topics	15:00:00	0:00:25	14:59:35	16:00:00	0:00:23	15:59:37
	A1.4 Populate Longlist (Term Document Matrix)	6:00:00	0:00:30	5:59:30	6:00:00	0:00:32	5:59:28
A2. Create Shortlist of prioritized Material Topics	A2.3 Gather Suggestions	2:00:00	0:00:01	1:59:59	2:00:00	0:00:01	1:59:59
	A2.1 Conduct Internal Consultation, A2.2 Conduct Voting Session, A2.4 Conduct Survey	NA	NA	NA	NA	NA	NA
	A2.5 Shorten Longlist by Results	5:00:00	0:04:00	4:56:00	7:00:00	0:04:00	6:56:00
A3. Test Outcomes (shortlist) against Stakeholder Opinions	A3.1 Analyze Existing Documentation, A3.2 Assign Importance by Stakeholder-category	20:00:00	0:00:35	19:59:25	22:00:00	0:00:37	21:59:23
	A3.3 Consult External Stakeholder Panel	NA	NA	NA	NA	NA	NA
A4. Conclude on Materiality (Final Material Topics)	A4.1 Conduct Presentation Session	NA	NA	NA	NA	NA	NA
	A4.2 Translate Results	10:00:00	0:00:01	9:59:59	4:00:00	0:00:01	3:59:59
	Sum	58:00:00	0:05:32	57:54:28	57:00:00	0:05:34	56:54:26

Table 2 – Actual Efficiency results

In Table 2, the numbers indicate time taken to perform the sub-activities corresponding to the eight pain points identified. The cells marked "NA" were not automated. The exercise conducted for data-set 1 took 58 hours manually. After

automation, the time was reduced to 5 minutes 32 seconds. One should however, not confuse this as the total time for the whole Materiality Assessment process, as there are few steps that have not been automated. The time accounted for in this illustration is only for the steps that we did automate, i.e., the pain points. This comparison is seen below.

Fig. 16 – Actual Efficiency results, Data Set 1
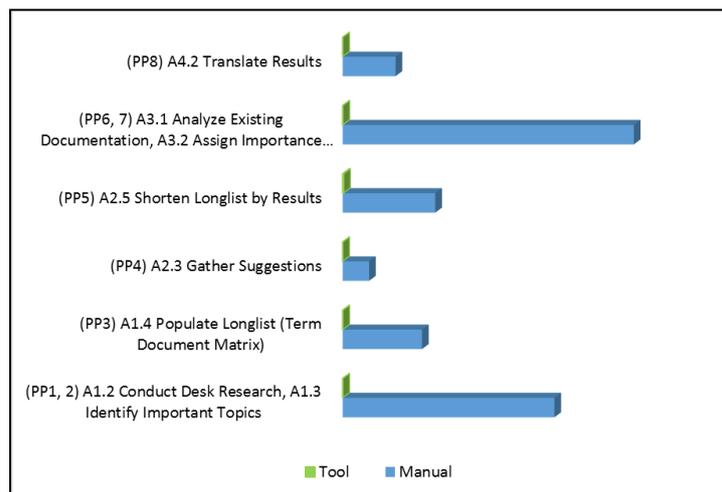
Fig. 17 – Actual Efficiency results, Data Set 2

The above illustrations depict that the proposed solution makes the process of Materiality Assessment faster. As opposed to the 58 hours of manual work, the tool requires less than 6 minutes, making the process 629 times faster for data set 1. For data set 2, the 57 hours is reduced to less than 6 minutes, making it 614 times faster. The tool proves to be particularly efficient for the document scanning tasks (pain points pertaining to conducting desktop research).

b.   Actual Effectiveness

The output from the tool (for each scoring scheme, and with varying thresholds) was compared against the manual output, and Precision and Recall metrics for each were calculated. First, the first 40 topics generated by the tool were compared the first 40 topics of Sustainalize for each scoring scheme. Then, for each scoring scheme, it was calculated how many positions a topic differs compared to the Sustainalize longlist. Consolidating the number of "matches" and "no matches" in the top 40, the precision & recall were calculated. The results of this are shown below:

	Count - longlist	Frequency - longlist	Relative [0,01] - longlist	Relative [0,03] - longlist	Relative [0,05] - longlist	Relative [0,07] - longlist	Relative [0,1] - longlist	Weighted - longlist
Number of topics in top 40 - Not matching Manual	20	23	21	20	21	23	23	24
Number of topics in top 40 - Matching Manual	20	17	19	20	19	17	17	16
Precision & Recall	50%	43%	48%	50%	48%	43%	43%	40%
Average difference in position top 40 compared to Manual	42	47	44	46	49	42	41	43
Average difference in position compared to Manual	32	33	32	36	37	35	36	32

Table 3 – Precision and Recall

Based on this exercise, it is seen that Frequency and Relative [0.03] yield the best results. However, this is based on one case study, and should be further re-affirmed before putting into actual practice. The precision and recall rates are in the 40-50% range. Suggestions for improvement of these scores are available in the "Conclusions and Future Work" section.

It is to be noted that the details of this exercise are not made available in the appendices, owing to the presence of confidential customer-data. These may be supplied on a need-to-know basis.

5.2   Perceived Parameters' Validation results

The survey designed to capture the perceived parameters, and circulated to a target audience of practitioners and researchers.

Limitations of the survey

The initial response rate was low (4 out of 10), and deemed insufficient for drawing a valid conclusion. We attempted to seek out the reason behind the low-response rate by speaking to the targeted audience of the survey. It was noted that the reason for low response rate was the limited interaction they had had with the tool. To remedy this, the tool was installed individually on the systems of the target-audience, and a familiarization period of two working weeks was designated. However, the responses were anonymous, and it was thus not possible to identify and address the non-respondents individually.

A demonstration of the tool was carried out at the European Green Office Summit, Utrecht. The audience consisted of novices and researchers in the materiality assessment arena. The tool was made available for use to the researchers, and the survey circulated to them. This expanded the target-audience to 13.

Post these remedial measures, we were able to increase the response rate from 40% to 61.5%. Owing to the time constraints of this research, it was decided to proceed with result-calculation with this response-rate.

Survey results

The results were exported to excel. Based on the approach suggested by De Vaus (2013), each variable was assigned a polarity of positive or negative based on the question. For example, the question "Overall, I found the tool difficult to use" was assigned negative, and "I found the tool easy to learn" was positive. The negative scores were converted to positive scores with the formula:

New Score = (Highest Possible Score + Lowest Possible Score) – Original Score

After the conversion, the scores of each respondent per variable were summed up. Then we proceeded to normalize these values, as the variables had different number of questions, and thus not comparable in the original form. The graphical illustration of these results is shown below. Details of the conversions and calculations are available in Appendix 5.
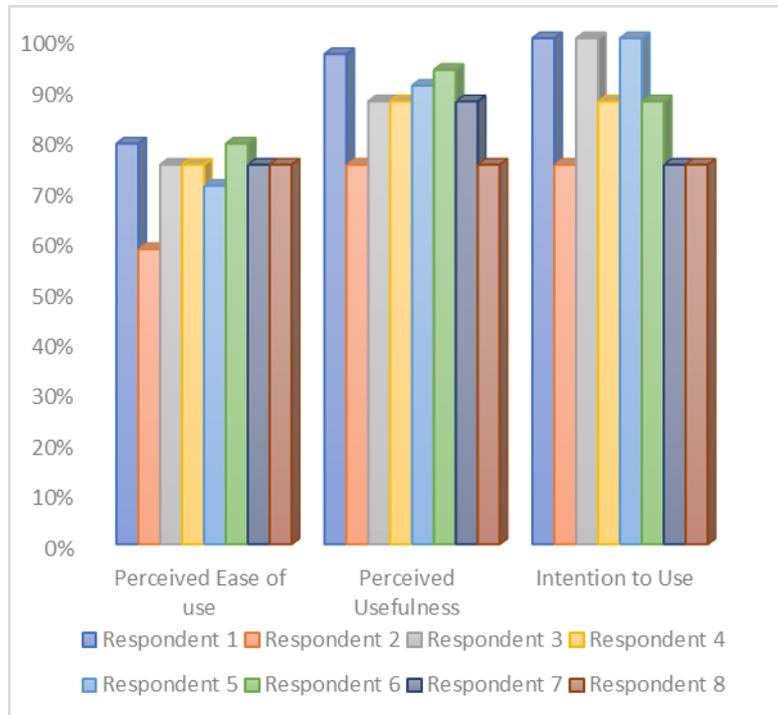
Fig. 18 – Normalized Scores of Survey Results

The average normalized score for Perceived Ease of Use was 73%, for Perceived Usefulness 87%, and Intention to Use 88%. The low score of the Perceived Ease of Use parameter was traced to the fact that the respondents had not received any formal training for the tool at the time of the survey, and were not familiar with the features and navigation. Although the tool had been made available for three weeks prior to the final results calculation, we propose that the validation exercise be carried out after the users have interacted with the tool on a daily basis for an extended period of time.

The Perceived Usefulness scored high, as the time saved by the tool compared to the current manual process, was immediately evident. This perception is backed up by the Actual Efficiency measures in section 5.1.

The Intention to Use, though received a high score, we could not measure the Actual Usage. This would depend on how familiar the respondents got with the tool. Discussions with the targeted-audience yielded that prolonged exposure and interaction with the tool could lead to usage of the tool in practice. Discussion with the project-owner yielded that the tool, with further enhancements, could be made commercially available to customers. At the time of the conversation, consultants were using it to speed up the materiality assessment processes internally, and for demonstrations to potential and existing customers.

6   Conclusions and future work

In this work, we have faced the dilemma of balancing rigor with relevance, and reaching the optimal trade-off. While keeping the necessary rigor in a master thesis, we have favored relevance, given the need to provide the best possible solution to the consultancy company involved. There was also the issue of disparities in the way the practitioners express their way of working (method in the book) and their actual practice (method in practice). We have tried to iron out these disparities to some extent by the use of formal models to capture the methods. Another perceived threat to validity was that the Case Study was conducted at one company, with two data sets. A larger population would provide more insight into the process variances.

The tool received favorable responses from the survey respondents and scored high on the efficiency parameters. However, the effectiveness parameters, Precision and Recall could be improved. We propose the following lines of future work for improving these:
- The longlist currently contains many abstract terms. For example, the tool would not find a perfect match for the phrase "Gender equality" in the documents. The resolution for this is to implement a learning module. The learning module would contain a Bag of Words feature, containing fringe terms, like male/ female/ sex ratio, which should be tagged with the longlist item of "Gender Equality". The algorithm would then be required to pose a question to the user on whether the word "male" or "female" is related to the Gender Equality phrase. Based on the user's response, this new learning would be added to the Bag of Words or the longlist.
- Another aspect that could be looked into is adding a thesaurus plug in. This would help the tool expand the bag of words by adding synonyms for each longlist term. The thesaurus and word packages available readily with Google may be used for this purpose.

The tool developed in this research may further be expanded to elicit data from sources other than the documents currently being used. This could include social media mining in order to elicit the sentiments of various stakeholder groups. We have compiled a template with examples of sources that may be helpful in this research. These sources include academic databases, commercial databases (open source and paid), as well as free flowing social media data. The template is provided in Appendix 6.

Another prospect that may be looked into is the usage of news on social media to identify upcoming topics in the respective sectors. However, one must be cautious in this exercise, due to the prevalence of fake news on social media websites. An algorithm has to be devised to filter out fake news, so as to avoid making misinformed strategic decisions. A suggested shortcut to this is using a website or services that provide tailor-made sentiment analysis services. However, this needs

further investigation and trials to find the best fit. The details of this investigation and suggestions are available in the Appendix 1.

More features could be added to the tool, which were part of the initial requirements, but could not be implemented due to time constraints of our research. There could be authorization segregation based on the role of the user, depending on how the company wants to implement it. It could be a good strategy to provide customers access to the visualizations only, while keeping the nuances of the matrix population as an asset for the consulting business.
A proposed architecture with these enhancement suggestions is shown below.
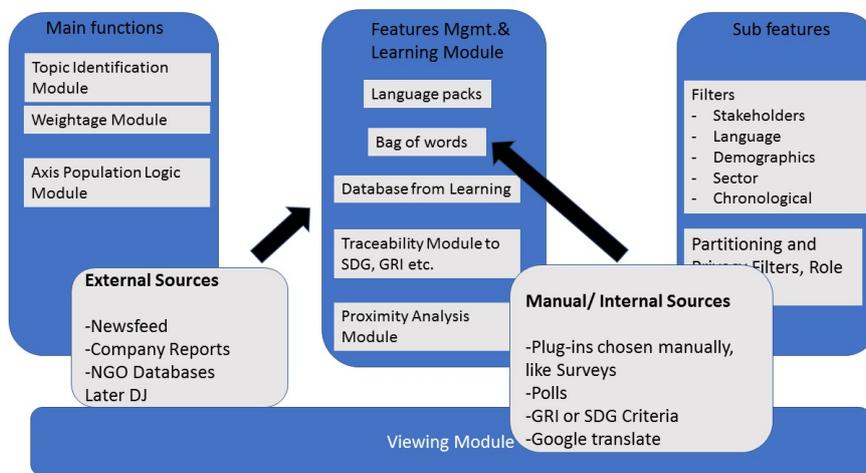
Fig. 19 – Proposed enhancement for the tool

The observed impact of the tool applied to the problem scenario is primarily the time-savings aspect. Commercially, these time-savings enable practitioners to eliminate or minimize time taken for redundant tasks, specifically, document scanning, term document matrix population, which takes up to 60 - 80% of the time in materiality assessment exercises. Automation of these tasks and the enhancements to the existing process make for a lucrative business-opportunity.
This tool is made freely available for academia, with the aim of helping researchers simplify their work with materiality assessment exercises. It could be applied by academic institutions to identify areas of concern and improvement, as noted in the Green Office Summit. Looking at the bigger picture of Socio-Environmental Auditing, this tool could serve as one of the facilitators to help auditors, NGOs and regulators to verify the topics requiring emphasis, or identified for reporting. The proposed future enhancements to the tool could be an interesting line of research for academics interested in the sustainability domain.

Acknowledgements

References

1. Agarwal, A., Xie, B., Vovsha, I., Rambow, O., & Passonneau, R. (2011, June). Sentiment analysis of twitter data. In Proceedings of the workshop on languages in social media (pp. 30-38). Association for Computational Linguistics.
2. Brunetti, G., & Golob, B. (2000). A feature-based approach towards an integrated product model including conceptual design information. Computer-Aided Design, 32(14), 877-887.
3. Calabrese, A., Costa, R., Levialdi, N., & Menichini, T. (2016). A fuzzy Analytic Hierarchy Process method to support materiality assessment in sustainability reporting. Journal of Cleaner Production, 121, 248-264.
4. Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. Management science, 35(8), 982-1003.
5. De Vaus, D. (2013). Surveys in social research. Routledge.
6. Eccles, R. G., & Saltzman, D. (2011). Achieving sustainability through integrated reporting. Stanf Soc Innov Rev Summer, 59.
7. Gray, R. (2000). Current developments and trends in social and environmental auditing, reporting and attestation: A review and comment. International journal of auditing, 4(3), 247-268.
8. Jain, T. I., & Nemade, D. (2010). Recognizing contextual polarity in phrase-level sentiment analysis. International Journal of Computer Applications IJCA, 7(5), 5-11.
9. Johns, R. (2010). Likert items and scales. Survey Question Bank: Methods Fact Sheet, 1, 1-11.

10. Kouloumpis, E., Wilson, T., & Moore, J. D. (2011). Twitter sentiment analysis: The good the bad and the omg!. Icwsm, 11(538-541), 164.

11. Manetti, G., & Becatti, L. (2009). Assurance services for sustainability reports: Standards and empirical evidence. Journal of Business Ethics, 87, 289-298.

12. Milne, M. J., & Adler, R. W. (1999). Exploring the reliability of social and environmental disclosures content analysis. Accounting, Auditing & Accountability Journal, 12(2), 237-256.

13. Moody, D. L. (2003). The method evaluation model: a theoretical model for validating information systems design methods. ECIS 2003 proceedings, 79.

14. Salzmann, O., Ionescu-Somers, A., & Steger, U. (2005). The business case for corporate sustainability:: literature review and research options. European Management Journal, 23(1), 27-36.

15. Roberts, J. S., Laughlin, J. E., & Wedell, D. H. (1999). Validity issues in the Likert and Thurstone approaches to attitude measurement. Educational and psychological measurement, 59(2), 211-233.

16. Singh, R. K., Murty, H. R., Gupta, S. K., & Dikshit, A. K. (2012). An overview of sustainability assessment methodologies. Ecological Indicators, 15(1), 281-299.

17. Ting, K. M. (2011). Precision and recall. In Encyclopedia of machine learning (pp. 781-781). Springer US.

18. van de Weerd, I., & Brinkkemper, S. (2009). Meta-modeling for situational analysis and design methods. In Handbook of research on modern systems analysis and design technologies and applications (pp. 35-54). IGI Global.

19. Vlaanderen, K., van Tuijl, G., Brinkkemper, S., & Jansen, S. (2013). Incremental Method Enactment for Computer Aided Software Engineering Tools. In Enterprise, Business-Process and Information Systems Modeling (pp. 370-384). Springer, Berlin, Heidelberg.

20. Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. International Journal of Research in Marketing, 27(3), 236-247.

Appendix 1 – Literature Review

The initial academic literature review was based on studying the background and context of the problem, to gain a better understanding of the underlying principles behind sustainability reporting. The learnings from this are summarized here.

The work by Milne & Adler (1999), though conducted almost two decades ago, gives meaningful insight to the emergence of environmental and social disclosure. At the time of the research, sustainability reporting was not considered mainstream, and was an up and coming issue, as predicted by experts. It delves into the reliability of the contents provided in disclosure reports, and highlight the need for a codification scheme requirement in this arena, precisely, "*To-date, no published studies appear to exist on the reliability of the coding instruments used for classifying organizations' social and environmental disclosures*"(Milne et al, 1999). This helped us validate the importance of this current research, and confirmed the benefit of the development of a Materiality Assessment framework. The discussion on "acceptable standards for social and environmental data reliability", is particularly interesting to our context, as reducing "subjectivity" in materiality assessment is one of the primary motivations behind our research. It identifies three types of reliability measures in content analysis – stability, reproducibility and accuracy. These measures inspired the validation criteria defined for our proposed problem-solution, by contributing to the features required in our envisioned final solution.

The experiment detailed in the paper, where content was analyzed by three coders, and the results compared to identify subjectivity, yields interesting insight to the underlying cause of the subjectivity issue. The reason behind the subjectivity is pointed out as "*because of various assumptions that are made about the unknown distribution of frequencies per content category*" (Milne et al, 1999)*;* this is an important part of the Materiality Assessment process, where experts codify the content based on frequencies of occurrence in text reports. Another similarity is that this codification, in our case, is also "categorized" manually based on sustainability framework standards. It is very aptly pointed out in this paper, that "*While it is not strictly possible to separate out the effects on coding performance of the coder from the coding instrument, it is possible to gain an understanding of the effects of each by holding one or the other constant….. by trialing a given method with coders who have different levels of experience and familiarity, it is possible to get an idea of the reliability of the method being trialed.*" (Milne et al, 1999) – which provides validity to our proposed solution being able to "control subjectivity involved due to human disparity."

The next work studied, *"An overview of sustainability assessment methodologies*" (Singh, Murty et al, 2012) looks into the various sustainability assessment methodologies currently in practice. It mentions "*The World Business Council for Sustainable Development (WBCSD, 1997), the Global Reporting Initiative (GRI, 2002a,b) and development of standards (OECD, 2002) were the key driver for adoption of sustainability management in industries'* – which led us to research these particular guidelines to gain more insight into the subject of the research. It highlights

the need for stable indicators, and their role in "*their ability to summarize, focus and condense the enormous complexity of our dynamic environment to a manageable amount of meaningful information*" (Singh, Murty et al, 2012). This explains the linkage of standards and frameworks like GRI to the reporting strategies of various organizations. It has excerpts from Lundin (2003), Berke and Manta (1999) and Spohn (2004), wherein it states the following about Sustainability Indicators,

"*SDIs can be used to: •   Assess and evaluate the performance. •   Provide trends on improvement as well as warning information on declining trend for the various dimensions of sustainability i.e. economic, environment and social aspects. •Provide information to decision makers to formulate strategies and communicate the achievements to the stakeholders.*". This was helpful in the "solution design part" of our research, wherein we seek to develop a tool for Sustainability Reporting. An interesting figure is obtained from this paper, "*The DPSIR framework for reporting on environmental issues. Source: Smeets and Weterings (1999)*", which helps one visualize the context of Sustainability Reporting.

"*The Driving Force Pressure State Impact Response (DPSIR) model is an extension of the PSR framework and has been adopted by the European Environmental Agency (EEA) and the European Statistical Office in 1997..... This framework defines the impact of human activities which exert 'pressures' on the environment and results in change in the quality and the quantity of environment conditions (the 'state'). Accordingly, society responds to these changes through environmental, economic and sectored policies (the 'societal response') for its adaptation (OECD, 1993). Response of society acts as a feedback to "pressure" segment through human activities (OECD, 1998)*" (Singh, Murty et al, 2012).

This gives us a very good understanding of where the stakeholders are placed in the arena of Sustainability Reporting. Having insight into the context helped develop further the requirements for the envisioned solution, based on what dimensions should be considered in sustainability reporting.

The next piece of work studied, "*Achieving Sustainability Through Integrated Reporting*", by Eccles and Saltzman (2011), because it brings to light "*the world's first guidance document for companies practicing integrated reporting*", which was a groundbreaking point in the arena of non-financial reporting. It identifies three classes of benefits of integrated reporting, "*internal benefits, including better internal resource allocation decisions, greater engagement with shareholders and other stakeholders, and lower reputational risk. The second is external market benefits, including meeting the needs of mainstream investors who want ESG information, appearing on sustainability indices, and ensuring that data vendors report accurate nonfinancial information on the company. The third is managing regulatory risk, including being prepared for a likely wave of global regulation, responding to requests from stock exchanges, and having a seat at the table as frameworks and standards are developed.*"

It pointed us to another valuable source, "*a free e-book, The Landscape of Integrated Reporting: Reflections and Next Steps, published in November 2010 following an integrated reporting workshop at Harvard Business School. Companies now have the benefit of an increasing body of literature on integrated reporting, and those writing*

about it have the benefit of an increasing number of companies that are practicing it." (Eccles and Saltzman, 2011).

Though the contents of this paper merely validated the already-mature motivation behind our research (and not any additional information), the main contribution of this paper was that it proved an excellent source for other literature, papers, regulatory documents, which we then studied for our research.

The next work studied, *"The business case for corporate sustainability: literature review and research options",* again did not yield much insight into new content, as much as redirecting to other new material. It lists out several theoretical and empirical studies on relationships between financial and non-financial reporting parameters, thus providing to yet again reaffirm the commercial motivation behind this research.

The secondary literature review was done with the aim of uncovering the technologies and trends in the arena of Materiality Assessment. The primary subject of this research was "how to identify material issues on social media, and how to interpret the opinions/sentiments of various stakeholder groups regarding these issues."

The first work studied in this context, *"Twitter sentiment analysis: The good the bad and the omg!"* (Kouloumpis et al, 2011), investigates *"Features such as automatic part-of-speech tags and resources such as sentiment lexicons have proved useful for sentiment analysis in other domains, but will they also prove useful for sentiment analysis in Twitter?".* It starts by defining the challenges encountered in analysis of Twitter data, such as the huge volume of topics covered, and the boundaries being blurry on what is reliable information and what is not. The work centers around building a lexicon based on which Twitter hashtags may be analyzed for positive or negative emotion. The research uses three different corpora of Tweets in experiments to analyse emotion in hashtags *("to identify positive, negative, and neutral tweets to use for training three-way sentiment classifiers",* Kouloumpis et al, 2011). However, the experiment conducted is quite elaborate and could not be repeated by us, and is hence not described here. The take-away for us from this paper was the affirmation of the hypothesis that Twitter data may be used to some extent to identify sentiments and opinions of stakeholder groups, keeping in mind the confidence-boundaries owing to the ambiguity in social-media jargon.

A very important part of our solution design emerged from the information in this paper. Instead of building a sentiment-analysis module from scratch, it redirected us to *"Companies such as Twitratr (twitrratr.com), tweetfeel (www.tweetfeel.com), and Social Mention (www.socialmention.com) are just a few who advertise Twitter sentiment analysis as one of their services".* It would have been redundant (and not at par quality-wise), for us to attempt and build something which is readily available for plugging into our envisioned solution.

The next piece of research studied was *"Sentiment analysis of twitter data"* (Agarwal, Xie et al, 2011), which focuses on building models for classifying Tweets as positive, negative or neutral. The experiment was based on three types of models, unigram model, a feature based model and a tree kernel based model, and at a later stage, permutations and combinations of these three models. This paper introduces us to two available resources, *"a hand annotated dictionary for emoticons that maps emoticons to their polarity and an acronym dictionary collected from the web with English translations of over 5000 frequently used acronyms".* This paper provides a simple

algorithm which could be adapted to serve as the backbone of our envisioned solution. Results indicate that this algorithm can analyse sentiments behind Tweets up to 80% accuracy. The figure below shows our interpretation of their algorithm, simplified to suit our needs:
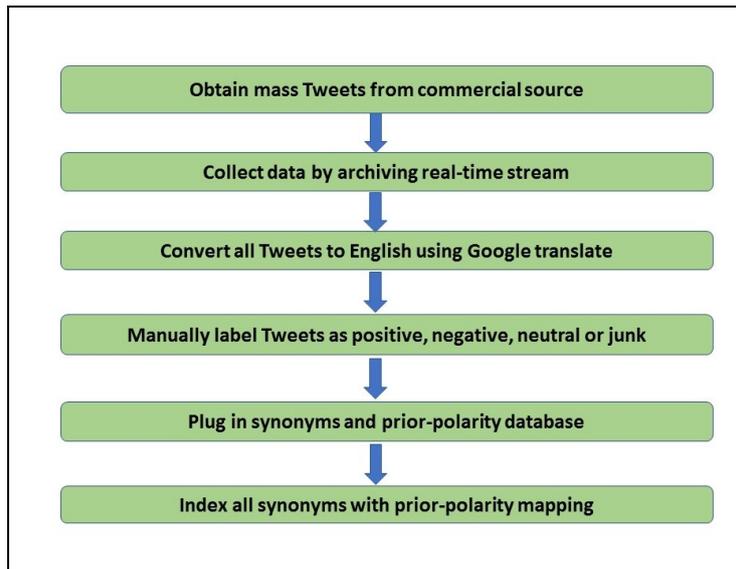
Fig. 20 – Algorithm for Twitter sentiment mining

The next work studied was "*Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis'* by Jain and Nemade (2010). This paper delves into the ambiguity in sentiment mining owing to demographic factors, and proposes the usage of different training data sets for various geographical locations, to achieve maximum accuracy. This is applicable to our proposed solution, as the issue of language interpretation and ambiguity can pose as a threat to the validity of output of our solution. However, this is not something that is in the scope of information sciences, and veers more towards a lexicology angle. Though the paper provides us useful insight as to the issues that might arise, it is too premature for us to take these into consideration right now. Once the prototype of the solution is at hand, we shall come back to this topic with the aim of refining our solution, if time permits.

Appendix 2 – Traceabilities and Modeling details

This section showcases the rationale behind the models created during this research, by linking them to evidences.

2.1 Mapping of Generic and Situational DPSIR Model

Each node in the DPSIR Model and their interactions are shown in the figure below.

Generic DPSIR Model	Explanation	Situational (Sustainability Reporting DPSIR Model)
Driving Forces	The factors that trigger the problem scenario.	• Awareness among customers & investors • Regulations & Laws • Peers & competitors
Pressure	The threat to the current state of affairs, created by the Driving Forces.	Demand for Sustainable & Ethical business, disclosure
State	The pressure threatens this situation or entity	• Current Business scenario (clients, market value) • Investors • Reputation
Impact	The threat could lead to these damages to the "state".	• Loss of Business • Negative image
Response	The Business should respond to the threats, in an attempt to resolve them/ minimize/avoid the damage, by means of these actions.	• Full Disclosure • Ethical Business

exert · applies to · causing · demanding

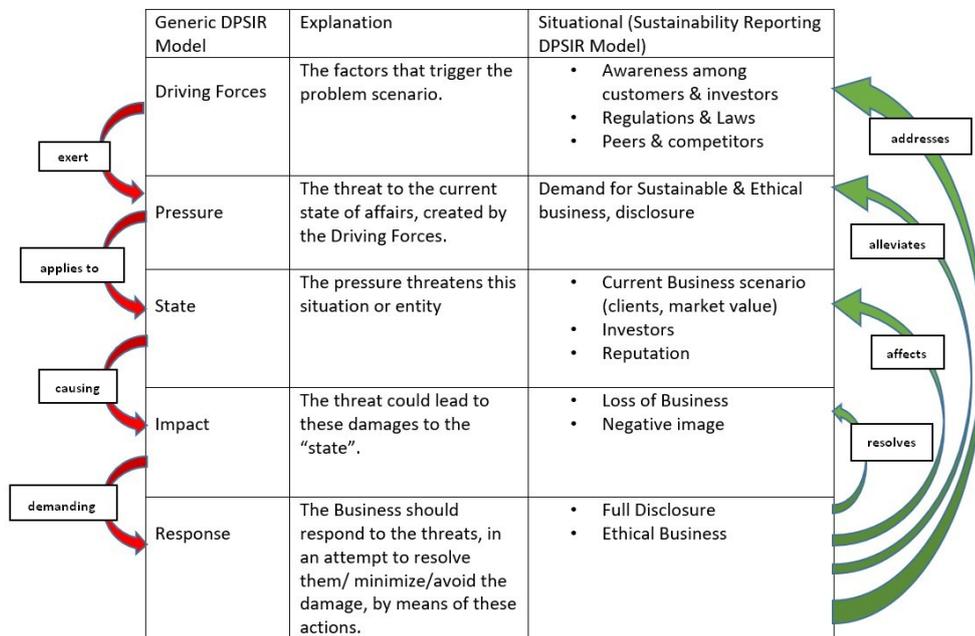addresses · alleviates · affects · resolves

Fig. 21 – DPSIR Model interactions

Traceability between Stakeholder groups in Stakeholder Model with documents reviewed during research

The table below traces the origin of the rationale behind the identification of stakeholder groups modeled in this paper.

Stakeholder Groups	Link Stakeholder Groups	Document type	Reference
Employees	Employees	Excel overview of results/data, PDF report	Delphi stakeholder questionnaire (DE)
		Proposition: Survey	Stakeholder Survey 2016
Consumers	Consumers	Mail	Consumer enquiries (UK)
		Mail	Consumer enquiries (DE)
		Word document	Consumer enquiries (FR)
		Excel overview of results/data, PDF report	Delphi stakeholder questionnaire (DE)
Supply Chain	Customers (Retailers) Suppliers	Word document, website	Tesco Sustainability Approach (UK)/ website
		Word document, website	Wholesale Sustainability Mission and Vision (UK)/ website
		Word document, website	M&S Plan A Summary on CSR Benchmark Selection (UK)/ website
		Website	Carrefour (http://corporate.carrefour.eu/Sustainable Development.cfm?lang=nl)
		Website	Auchan Holding (http://corporate.carrefour.eu/Sustainable Development.cfm?lang=nl)
		Website	Ahold (https://www.aholddelhaize.com/en/about-
		Website	Ekoplaza (https://www.ekoplaza.nl/pagina/samen-
		PDF report	Waitrose Way Supplier Awards 2016
		Excel overview of results/data, PDF report	Delphi stakeholder questionnaire (DE)
		Proposition: Survey	Stakeholder Survey 2016
Investors	Shareholders Investors and banks	PDF report	MSCI ESG Controversies report
		PDF report	MSCI ESG Ratings report
		PDF report	Sustainalytics ESG Report
General public	Society	Excel overview of results/data, PDF report	Delphi stakeholder questionnaire (DE)
	Governments	Newsarticles, websearch	News media analysis
	Experts	PDF report	Food Watch Report 2014 - Lost in the
	Media	Website	VBDO (http://www.vbdo.nl/)
	NGOs	Website	Solagro (https://solagro.com)
		Website	CDP (https://www.cdp.net/en)

Table 4 – Traceability of Stakeholder Model

2.2   Details of Process Deliverable Diagrams

This sub-section contains the descriptions of the activities, and definitions of the concepts found in the PDDs. The activities from the Design Cycle PDD are described below:

Activity	Sub-Activity(s)	Description
Problem investigation	Literature study	Study of academic literature to gain a deeper understanding of the problem.
	Case study	Observation, Apprenticeship and investigation of the problem scenario and processes at the identified company.
Treatment Design	Requirements Elicitation	Based on the above sub-activity, gathering of the requirements to resolve the problematic phenomena, and refining them after analysis.
	Solution Design and Development	Conceptualization and design of the potential solution (tool).
	Solution Implementation	Demonstrating the solution to the stakeholders, in order to improve and iteratively develop till the final solution is reached.
Treatment Validation	Apply Validation Criteria	Checking if the proposed solution fulfills all the requirements and resolves the identified problems. This was done by means of pre-defined metrics and conducting a survey amongst potential users of the tool.
	Assess Validation Results	Analysing the survey results and metrics values to arrive at a conclusion on the tool being a valid solution.

Table 5 – Activity table for Design Cycle PDD

The deliverables of each of the activities described above are defined in the table of concepts below.

36

Concept	Type	Definition
LITERATURE STUDY REPORT	Standard	A documentation of study of the academic documentation, to further verify the existence of the problem, and uncover existing research, based on data from Google Scholar and other academic databases like DBLP.
CASE STUDY REPORT	Open	A documentation of the investigation of the problem scenario at the selected company, consisting of field observations, apprenticeship, and data collection through interviews.
REQUIREMENTS DOCUMENTATION	Open	A documentation of the requirements for the proposed solution that would resolve the problem scenario/ phenomena.
METHOD SPECIFICATION	Open	The formal representation of the Materiality Assessment method, in the form of Process Deliverable Diagrams and other models.
BUSINESS CASE	Open	A dcoumentation of the results of the feasibility study of the commercial viability of the proposed solution, using SWOT Analysis.
TOOL PROTOTYPE	Open	First working version of the solution.
VALIDATION PLAN	Standard	A documentation of the criteria that would deem the solution valid for the problem, by employing metrics and survey.
VALIDATION RESULTS	Standard	A documentation of the evidence of how the solution fares against the VALIDATION CRITERIA, and analysis of the validation results.
PRESENTATION TO STAKEHOLDERS	Standard	A closing meeting summarizing the research, and demonstrating the solution.

Table 6 – Table of Concepts for Design Cycle PDD

The activities in the PDDs describing the Materiality Assessment Process (the high-level PDD, the detailed process PDD, and the improved process PDD) are described below, followed by the Concepts.

Identifier	Activity	Sub-Activity(s)	Description
A1	Create Long List of Material Topics	A1.1 Collect Relevant Documents	Gather the input documents for Materiality Assessment from relevant frameworks, sector and stakeholders.
		A1.2 Conduct Desk Research	Peruse all the input documents.
		A1.3 Identify Important Topics	Identify material topics matching with Long List based on the document perusal.
		A1.35 Select scoring scheme	This step has been introduced as a process-enhancement by the tool, and did not exist previously. The user may select one from four scoring options available in the tool, and based on this the Long List would be populated.
		A1.4 Populate Longlist (Term Document Matrix)	The act of documenting the matches found in the previous step in the Long List. Based on the occurerences and scoring scheme, each item in the long list is assigned a score.
A2	Create Shortlist of prioritized Material Topics	A2.1 Conduct Internal Consultation	Meeting with the internal stakeholders of the company (for which the materiality assessment is being undertaken).
		A2.2 Conduct Voting Session	For the highest scoring Long List items (based on a pre-defined cut-off), obtain the internal stakeholders' opinion on which topcs should make it to the short-list.
		A2.3 Gather Suggestions	Documenting the results of the previous step.
		A2.4 Conduct Survey	Send out survey to other internal stakeholders of the company, to take their opinion into consideration.
		A2.5 Shorten Longlist by Results	Incorporate the results of the survey and voting into the Long List, thereby selcting a shorter list of topics.
A3	Test Outcomes (shortlist) against Stakeholder Opinions	A3.1 Analyze Existing Documentation	External stakeholder documents are perused.
		A3.2 Assign Importance by Stakeholder-category	Each topic in the short list is assigned a weightage of importance based on assessment of external stakeholder documents.
		A3.3 Consult External Stakeholder Panel	Affirm with stakeholder panel that they agree with the output of A3.2.
A4	Conclude on Materiality (Final Material Topics)	A4.1 Conduct Presentation Session	Create and present the results of all the previous steps and final output at the company.
		A4.2 Translate Results	Plot the identified list of material topics from the exercise on to a the materiality matrix (graph).

Table 7 –   Activity table for Materiality Assessment Process PDDs

Concept	Type	Definition
SUSTAINABILITY FRAMEWORKS	Standard	Globally accepted guiding principles for sustainability reporting, to help organizations identify what topics to report on in their disclosure reports.
COMPANY STRATEGY AND APPROACH	Standard	The guiding principles and business values defined by and adhered to by the organization in question.
PEER REPORTS	Standard	Sustainability disclosure reports of competitors of the organization in question.
BEST PRACTICES	Standard	Collection of industry best practices accumulated based on professional/commercial experience, deemed as being most effective.
TACIT SECTOR KNOWLEDGE	Standard	Knowledge possessed by practitioners, based on field experience and domain expertise. This knowledge may or may not be formally documented.
MEDIA ANALYSIS REPORTS	Standard	Data collected from media channels, with the aim of providing insight into patterns and trends regarding a particular subject.
PREVIOUS REPORTS	Standard	Disclosure reports of the organization at hand, dating back to a specified period of time.
LONG LIST	Open with attributes	Exhaustive list of sustainability topics, serving as the basis for the final list of material topics to be disclosed. In the existing process, it was observed to possess attributes of categories, items and occurrence. In the improved process, an attribute of score was added.
WORD CLOUD	Standard	A visualization of words in cloud form. In our case, there is a word cloud based on prevalence of words in the input documents, and another word cloud based on matched occurrences between input documents and LONG LIST.
PRIORITIZED LIST OF TOPICS	Standard	The list of possible material topics sorted based on priorities assigned to them by internal stakeholders.
NEW TOPICS LIST	Standard	The shortened version of the prioritized list of topics, shortened based on a pre-defined cut-off (for number of topics to be included).
OPINION ON LONG LIST TOPICS	Standard	The documentation of the stakeholders' opinions on the items proposed to be included and excluded for the SHORT LIST. This is obtained by a voting session.
SURVEY RESULTS	Standard	The results of the second round of opinon gathering from fringe stakeholders to re-affirm the selcetion of topics.
SHORT LIST	Open	The shortened version of the prioritized list of topics, arrived at based on the scores and priorities assigned to the LONG LIST ITEMS.
FINAL VALIDATED SHORTLIST	Open	A refined version of the previously shortened list, based on the results of the survey and opinion gathering voting sessions.
MATERIALITY MATRIX	Open with attributes	A graphical representation of the final set of material topics selected for sustainability reporting. The plotting is done based on importance assigned to the topics by stakeholder groups.

Table 8 – Table of Concepts for Materiality Assessment Process PDD

Appendix 3 – Case Study Transcripts

This Appendix is divided into three parts. In the first part, we have documented the Materiality Assessment Process followed at Sustainalize, Utrecht. This documentation was based on internal company documents, and the insights obtained during the interviews with the consultants. The second part contains the interview transcripts. The interviews were geared towards getting insights into the actual process, and hence they were more open-ended and discussion oriented (rather than with a fixed set of questions). The third part, is the researcher's analysis and interpretation of the interview data and apprenticeship.

3.1 The Process

This section is meant to give the reader first-hand insight into the workings of the process, and the decision-making tribulations faced by the researchers, and is hence written in an informal way.

The steps below present an elaboration of the method specification.
Step 1: From long-list to short list
The goal of this step is to arrive at a short list of relevant topics, which in a later stage can be presented to both internal (step 2) and external stakeholders (step 3) for further assessment and ranking.

The first step of their approach is to generate a long list of sustainability topics and convert that objectively into a short list of topics. A comprehensive list of topics is the first key step in order to ensure that all relevant dimensions are included in the assessment and that no bias will occur. This legitimizes the outcomes of the materiality assessment – and the choices made in the process – towards internal and external stakeholders. The selection of topics will therefore include topics that do not (yet) have a visible link with the current business practices but may become urgent in the coming years.

In order to arrive at a long list of topics they will perform desk research using at least the following sources:

- The current topics covered in COMPANY X's strategy and –approach;
- Sustainability frameworks such as:
  - GRI (including sector supplements);
  - SASB;
  - ISO26000;
  - UN Global Compact;
  - Carbon Disclosure Project/Dow Jones Sustainability Index;
  - FTSE4GOOD;
- Sustainability reporting by peers;
- Sustainability reporting by external stakeholders such as: suppliers, business partners, customers, etc.;
- Best practices in sustainability reporting;

- Their sector knowledge and trends and developments in the sector;
- Topics covered in the media.

They will present their long list of topics in a brief session with COMPANY X.

After they have assured that all relevant topics are included in their assessment and that no angles and topics are missed, they will objectively rank the issues in order to arrive at a short list. This is important in order to have a manageable number of topics which in a later stage (step 2 and 3) can be presented to both internal and external stakeholders for further assessment and ranking.

They have proven methods and tools available that facilitate them in arriving at a practical and realistic list of topics. The following dimensions will be taken into account during this process:
- topics that are dominant in (reporting) frameworks;
- topics that are dominant in external reporting by peers;
- topics that receive high media coverage;
- topics that are included in the current strategy;
- topics that COMPANY X wants to test among its stakeholders;

Knowing that decisions taken in this stage might have a profound impact later on in the process, they will discuss the outcomes in an interactive session with COMPANY X. This will enable them to arrive at a realistic and manageable set of topics for inclusion in the internal and external stakeholder dialogue.

Step 2: Internal consultation

To prioritize the topics from the short list they propose to start with an internal consultation. They propose to do so in an interactive voting session in which the shortlist from step 1 is discussed in detail. The session is aimed at assigning importance to each of the topics included in the shortlist from the viewpoint of COMPANY X International. They use an interactive voting system allowing everybody to cast its vote anonymously. Results will be translated in graphs real time and can be used for valuable discussions. They will prepare this session and facilitate in its execution. After the session they will analyse and summarize the results as input for the final materiality assessment and matrix (see step 4).

During this process they are open to suggestions to further expand or shorten the selection of topics. Practical and realistic outcomes are important during this process while meeting G4 requirements. To have an effective and efficient process they suggest limiting this working session to a select number of key players only (max 10/12). Experience shows that this yields the best results.

Additionally, they will send out a brief survey to existing internal networks such as the PA-network and the SHE-network. They will summarize the findings from both exercises.

Step 3: Test outcomes against external stakeholder opinions

To complete the process of materiality assessment – and comply to the G4-guidelines - they need to test the outcomes of the preceding steps again against the opinions of external stakeholders. They propose to do so based on an analysis of existing documentation. Instead of requesting additional time from stakeholders, this approach therefore builds forth on existing sources resulting in both time and cost efficiencies.

Organisations communicate on a regular basis with their stakeholders. By reviewing the documentation of these engagements, they can model the importance stakeholders assign to key sustainability themes (shortlist from step 1 and 2). This approach is in line with the minimum requirements as set out by the G4-guidelines.

They support the process of determining and requesting the desired documentations within the organization. After gathering all relevant sources, they will analyse the information. Based on this analysis the outcomes of the shortlist of step 1 and 2 will be ranked according to the perspective of the external stakeholders. They will categorize the outcomes per main stakeholder group (investors, employees, government, suppliers, etc.) to gain insight in the relative weight that different stakeholders place on topics.

To finalize this step, they propose to test the outcome with the external stakeholder panel. They understand from COMPANY X that these panel is soon to be established to institutionalize the outside-in view. They will prepare and facilitate this validation (either by conducting a review round by mail or by means of an interactive session). A sample of output of this step is shown below.

Topics based on prioritisation of internal & external score			
2016 Shortlist topics	Internal score	External score	Averag
1 CO_2 emissions and climate change	3.1	3.6	3.38
2 Regulatory compliance and labelling	3.9	2.9	3.37
3 Healthy food	3.6	3.1	3.33
4 Consumer health and safety	4.0	2.6	3.30
5 Organic	3.3	3.2	3.22
6 Ethics in the supply chain	3.3	3.0	3.16
7 Origin of raw materials used	3.3	2.7	3.01
8 Food waste management	3.4	2.6	3.01
9 Fair working conditions	3.6	2.3	2.94
10 Healthy lifestyle	3.1	2.7	2.94
11 Fair trade	3.1	2.7	2.90
12 Genetically-Modified Organism (GMO)	3.6	2.2	2.88
13 Green entrepreneurial culture	3.7	1.8	2.75
14 Packaging	2.7	2.4	2.56
15 Sustainable employability	3.3	1.8	2.53
16 Educating consumers about sustainable food	3.1	1.9	2.51
17 Community involvement	2.6	2.3	2.41
18 Fair and ethical advertising	2.7	2.1	2.40
19 Biodiversity	2.4	2.0	2.21
20 Use of (sustainable) palm oil	2.6	1.8	2.17
21 Water efficiency	2.0	2.1	2.04

Fig 22 – Sample output

Step 4: Conclude on materiality
The final step in their support is to combine the outcomes from all previous steps and translate this into a final overview of material topics. Through an interactive session they will elaborate on their findings.

They translate the results of this session in a 'materiality matrix' to finalize the materiality assessment and to reflect on the contents of the sustainability report. They

will do so per significant stakeholder group. A 'materiality matrix' is a schematic representation of the importance of topics, both from the
perspective of COMPANY X International as well as its key stakeholders. As such, it prioritizes topics. The matrix is
ideally suited for inclusion in the sustainability report as it legitimizes the topics reported on. They will also assist in the link to applicable GRI-aspects.

1.1 Interview & Discussion Notes

This section talks about the process, problem and solution requirements, as discussed with the experts at Sustainalize. It is meant to give the reader first-hand insight into the workings of the process, and the decision-making tribulations faced by the researcher while undertaking the materiality assessment exercise, and is hence written in an informal way. The requirements elicitation was done on the premises of Sustainalize Utrecht, by means of interviews with the in-house experts in the Materiality Assessment domain. Some interviews are conducted with their partner company F19. It should be noted that the information mentioned herein was captured at preliminary stages of the project, and not all were translated into functionalities eventually.

Justus, 8.3.2017
Marcel & Nick have discussed with DJ, DJ has a database
We can use their existing db, using our software
Dj sells license for people who want to use their database
Several approaches:
o Develop own software for data mining
  Or
o Use existing tools & databases
  Adv

    → Having a name like DJ involved will help sell tool more
  Disadv
    → This license is too expensive (20K)
    → Copyright issues
    → Can more than one customer use this or separate licensing fees for each customer
    → Details + complexity unknown as of now

So, they have spoken to another company called Meltwater https://www.meltwater.com/nl/
  o Meltwater provides similar services as DJ, but lower price

The goal of MA – fair representation of what stakeholders think
Categories of stakeholders: {classical}
    → employees
    → Investors
    → Customers
    → Suppliers

➔ Authorities / legal & regulatory bodies

Social media can be analyzed to get these requirements, e.g., klm has text mining software to see how their name pops up, and they reply to +ve as well as -ve stuff to address concerns of customers

This is powerful, but it addresses only 1 stakeholder group, i.e., the client

There are 3 approaches/ models for Sustainalize MA right now:
1. Round table
   ➔ With key stakeholders
   ➔ Structured dialog
   ➔ Good way to understand exact demand
   ➔ Drawback – is the representavity sufficient and accurate?   Sample size vs population size
   ➔ Might work for some stakeholder groups, say for example, authorities (contact point can cover concerns of those whom he represents). This is ok to some extent for investors as well. But customers have diverse requirements

2. Desktop analysis
   Most companies are already in contact with stakeholders, (obv. As a company you are always in touch with your customers), but this might not always be structured. This communication/meetings can be translated into requirements if properly analysed etc. and may be documented to approximate the stakeholders' expectations groupwise. Also, Market Analysis Reports can be used in the same way!

3. Survey
   Send the shortlist of topics to a large number of stakeholders, rank them from our perspective and ask them to grade them from their perspective.
   Stakeholders are weighted (some are more important than others)
   Adv ➔ Broader reach, more representation, better sample/population ratio
   Disadv ➔ It is subjective, based on perception only!
   Long list ➔ precheck based on desktop analysis ➔ find predominant topics ➔ shortlist (include these in survey) ➔ ask stakeholders to rank each of the topics ➔ generate output per stakeholder group (weighted)
   See page 32 in guidelines of GRI G4, see diagram in this page. It is very generic, fits all three above approaches. But this is what the GRI wants to see. Cornerstone of this is materiality assessment.

Financial reporting	Sustainability reporting
Scoped for your own organization	Not only your own organization, but upstream and downstream (e.g., suppliers and consumers, whole value chain)
You have financial/operational control over what is in the scope of reporting	You don't have financial/operational control over what is in the scope of reporting
Objective – human independent	Subjective – human dependent
There is a clear framework for the reporting ➔ There are standards for reporting ➔ Various orgs reports can be compared because there are standard parameters like profit, revenue, etc.	There is no framework ➔ There are some, like GRI etc. but who's to say who is following what? ➔ Scope is too broad, no boundaries

Table 9 – Background of MA

Some notes about Unilever example (Value Chain video):
- ➔ They can use renewable energy or reduce footprint etc. in the factories but that will not have so much impact
- ➔ The impact will be much larger if they make a product that incorporates sustainability, for example, a detergent that does not need hot water.
- ➔ If financial, they report on their own factories only, but if value chain, they report upstream and downstream too
- ➔ Topics are different, scope is fundamentally different
- ➔ Impact is outside organizational boundaries
- ➔ Sometimes CSR agenda will be linked to future risk management, like unilever doesn't want to destroy the env cz then they can't grow cacao beans and produce chocolate for which there is a huge growing market

Having inward focus can make a company obsolete e.g. nokia. You need to look outside rather than inside, to keep up with the changing times
CSR is long term, very limited short-term incentive for CSR
CSR is niche
95% of focus of any company will be $$$
But nowadays, the remaining 5% is being given a lot of importance

       ➔   Precompetitive
       ➔   Managing future risk ➔ STRATEGIC RISK MANAGEMENT
       ➔   Intrinsic motivation (?)

- Who knows what topics will be of importance after 10 years? For example, child labor wasn't a big focus say 20 years ago but now all brouhaha
- In the MA of unilever, profit might be on top right
- But this will be for most commercial orgs, because they want profit! Cannot run a company without profit, it is of course the most pivotal.
- However, MA should include economic as well as non- economic indicators.

Alissa, 14032017
Basically, for all tasks:
- Support creates,
- Manager checks & does customer communication
- Partners get involved only if required, for very important stuff

Workshops in the internal consultation phase. the presentation is shared during the workshop
Add to internal prioritization
Test outcomes – depends on how tested
- Survey – basis is shortlisted (approved)
  Made by support, checked by manager, approved by customer
- Round table – Nick & Justus
  Created by support
- Desktop analysis – support + manager

Once we have the results, the support creates mat mat, and full presentation, excel with all calculations, ppt with details of gri stuff and links, etc.

Sometimes:
Describe every step, make memo for accountant (Del, PWC)
Show every choice made, justify all inclusions/ exclusions, share all rationale, etc.

In April, presentation by Alissa ➔ different approach, to make it more standardized
The HOW part will change, not the WHAT parts. More like fine-tuning the rationale, also long list/ short list creation, revisiting the hygiene topics, will be in the long list, but not in the short list.

Analysis done lately, focus on topic, not on the categories.

Does long list have all the topics?
If yes cool, if not revisit

Actual analysis, put x in the matrix if topic is in Annual report and strategy of the company

For customers' peers -   only look at strategy, don't put generic stuff in there, focus more on the company's USPs, (only if it is very important)

Sometimes also include media analysis in long list/ short list → news mining

X s counted in score, filter – how many topics are there

Try to include 40/50 topics before shortlist

Select and put in shortlist

Then filter – "all reported by vested (example)

So as not to miss out on anything important to customer

Exclude all topics ONLY mentioned in frameworks

Why?

- Not always the same in practice, to rigorous
- If peers are not reporting, it is maybe not that interesting!

Then make clustered shortlist → special for this customer, with their terminology (for e.g., from annual report)

Part of shortlist, specialize according to strategy and feedback

Once manager is satisfied, support can start working on the definitions

Willem does this – clustering all definitions used in the past few years, but still must tailor to context, else too generic

Based on annual report of customer + past work done

It should be aligned to the business of that particular customer, context specific

BOM example

Session – voting system, internal scores

Workshop – show presentation, explain why mat mat, what are the steps in mat mat

Why do we have this, why do we need this, how does it help you, trace back to strategy of company and annual reports

Show results at end of the session

Survey conducted at vested

Intro, explanations, background, context, then ask them to rank

1   To 10 scale Alissa prefers,

Try to have even, so they can't give a neutral answer, have to make a decision!

Provide definitions along with the survey

After rating all, ask them to RANK top 5

- This has been done only for klm till now
- They rank individual topics, NOT categories
- Based on ranking, additional 25% weightage if named max. times
- Others get less weightage depending on how many mentions
- Based on these percentages + the survey results create actual score for the materiality mat
- Score + additional % if in top 5 ranked
- (score*%) = final score

- Cannot compare how mm s of different companies are made, but can compare the outcomes
- Exact score not imp
- Relative performance is imp
- Best practice that Alissa is doing:
- Line for averages, so see top right quadrant → more important
- Then decide which topics are more material, to include them in report
- Sustainalize will suggest, but it is ultimately the customers call
- Explain if including something that is not in top list
- Topic → GRI 4 indicator → explanation →2017 GRIS
- Not always done, depends on proposal (doing for KLM because they have a lot of hours). Going to be doing more in the future.

3.3   Requirements Analysis

The researcher's understanding of the problem led to the following insights. First, the relevant topics need to be identified from the social media data, based on occurrence. These topics will differ based on the industry sector. A software giant may find workplace ethics and opportunities more relevant, whereas a retail goods manufacturer might find it relevant to look at human rights, labor rights and supply-chain aspects. Second, the sentiments of the around these topics need to be identified. People voice their concerns via Twitter, and these posts are seen and shared and have a strong influence on the perceptions of the stakeholders. Are stakeholders concerned about any particular topic more than other topics? Is any particular theme trending in the specific sector?   For example, an airline company wants to conduct a disclosure exercise (the practice of making known to the general public how a company is faring on the Material topics, as seen in the example Nestle report in the previous section). It would be helpful to know what topics they should focus on in their reports. Once the topics are identified, for example, "emissions", it would help to understand how their stakeholders feel about emissions. Is it a concern for them generally? How the airline company handles emissions and related disclosure could affect potential customers. In addition to the identification of these afore-mentioned topics, this whole process could benefit from having visibility on what issues are occurring related to these topics (for example, auto manufacturers concerned about emissions should be aware of news like the Volkswagen emission scandal).

The proposed solution has been assigned a name "The tool", for easy referencing.

<u>Functional requirements:</u>
1. The tool should identify material topics based on the sector and region in which a particular company(customer) operates.

2. It should assign weightage to these identified topics, based on which a customer would decide what to report on.

- KPIs – It should be possible to see the performance of the company based on their chosen KPIs related to the afore-mentioned Material topics. The KPIs shall be decided by Sustainalize/ customer/ based on GRI indicators. We need to define parameters that would capture these KPIs.
- We need to capture these parameters for each company over time so that thy may be able to compare their performance over time. One day if we suddenly want to see KPIs, we cannot do this out of the blue, because the KPIs need to have been configured at the beginning, and the parameters consistently captured over a period of time.
- We should be able to show trends in their KPIs vs. the KPIs of their Peers? Or is this confidential? Atleast we can show them their performance vs. average performance in their sector? Some sort of **KPI dashboard**
- To be able to capture the KPIs, customers may need to plug in their supply-chain data (As in, the figures). Does Dow Jones already have this? – Ask Nick!

3. Source data from sources in Source Document -> standardize (AIM is to auto-populate the Y axis)

4. Core should be in English, but then we should have the option of choosing region/language/country
    - Language interoperability (English, Dutch, French, Spanish, German, Mandarin/ Cantonese)?
    - google translate plug in good enough?
    - Chinese media restrictions need to be taken into account. Maybe local source is better there?

5. Interactive interface, customizable according to Customer/stakeholder's preferences:
    - Do they want to plug in Survey?
    - Input from Round table -> Can pull data directly from voting system? (we don't want to completely eliminate the Round table or face to face setting, because this induces engagement factor, and the dialog has value from our point of view)

6. The tool should accommodate the following **<u>Dimensions:</u>**
    - Demographics -> Region specific
    - Stakeholder groups
    - Different approach for each stakeholder group?
    - Internal vs. External Axis
    - Sector Specific

7. What about prediction/ forecasting?

8. Traceability 1 – To source - Can we have some sort of traceability, to weed out one particular source? We need to maintain traceability so that if we see social media (e.g., Facebook) is repeatedly giving false indicators, we need to scrap this source. So, each data item needs to be tagged to a source! -> How does GDRC do this? Check their Serendipitous tagging system!

9. Traceability 2 –
   - **Topics** should be mapped to GRI aspects
   - **KPIs** should be mapped to GRI stuff as well.
   - There should be **search** based on GRI criteria

   – only GRI? Or do we need more like SDG etc.? Maybe for prototype, GRI is sufficient. More can be added later. This will require updating on our part every time GRI updates their standards, it should be factored in as a recurring time-investment.

10. Define Search function – whatever is traceable is searchable.

Non-functional requirements:
11. Performance? How many Customers can be supported in parallel in the busiest of times? – parallel usage is not going to be an issue, according to Justus. Not initially, atleast.

12. Location - On premise/ Cloud? -> discuss with Eindhoven
    - Customer concern - is the server going to be in EU?
    - EU regulations – check + Marcel knows more (servers for EU customers should be in EU)

13. Different log-in for each customer? Or admin does all?

Product Diversification:

   - Option1 – We provide it as a **service/consultancy** – 1 off check – Customer tells us, hey can you run the check for me this weekend? We would run it, and have our internal expert check it and then get back to them.
   - Option 2 – We provide it as a **product** – They can check by themselves and continuously track their performance. (For this we need the tool to show indicators/KPIs).

14. Privacy - How to partition the backend data or the sources? We don't want one Customer to be able to run an analysis with another Customer's data?

- It would be nice to build up on sector specific data, use the sector knowledge to advise customers. For example, if we are doing the exercise for Heineken and collect lot of data for breweries (from survey or web?? – what did Justus mean), it would be a waste to not use this data for other breweries.
- However, if we are plugging in company specific survey data, then we need to make sure it's not traceable. We need to make it anonymous.
- Find a balance between the above two.
- See the point on KPIs. Should customers be able to see others' performances? What if it is anonymized and shown by sector?

15. What level of access will customer have?
    - each to have own log-in
    - Sustainalize to have Master log-in
    - If customers are plugging in their supply chain data, for the KPIs, and do not want to share, then their stuff will have to be partitioned
    - We need their permission(?) for using their specific data in the sector-performance snapshots!

16. Speed (depending on where the data source is):
    - Do we fetch the data and populate our database already?
    - Or is this done real-time? – Near real time is good, but this will not be an issue when we market it as a service. If someone is buying it as a product and will use it independently without our intervention, then this is very important.

17. How to ensure quality/ reliability? (actually, a problem related to the source) Do we provide confidence levels or how? Even though this would be useful for us internally, for the Customer, they don't want any uncertainty. They should have the assurance of reliability if they are to pay for this the tool!

18. Accuracy of results – (In the beginning only, perhaps) – After the analysis to capture material topics is run, we would want to have someone at Sustainalize review it.
    But does this reduce the benefit of the tool? Whole point was to automate it and take less time, or have the customers do it themselves? (See diversification for how Justus proposes to tackle this), should ultimately auto-populate the Y axis of the Materiality Matrix (External stakeholder axis).

Issues:
1. Sometimes one news item gets circulated too many times, making it seem important in the moment. We need to not make impulsive decisions based on such stuff! Because strategy is long term, and cannot be decided by one flashy

news item that might die down tomorrow. How to distinguish between something that will become a trend/issue, vs. something that will die down tomorrow?

Plug in company internal data for x axis risk analysis so that we can have indication of how issues trend

Should have timeline settings like company can choose if they want to see trending stuff for past 1 year or 1 week

2. How to distinguish false news when it comes to social media data?

3. What does NSA do? Samsung's report with the third axis?! If we have three axes – z axis showing what is the probability of this issue…. Too complex for first prototype, but should be looked at later for sure.

Based on the above functional requirements,   was perceived as shown below:
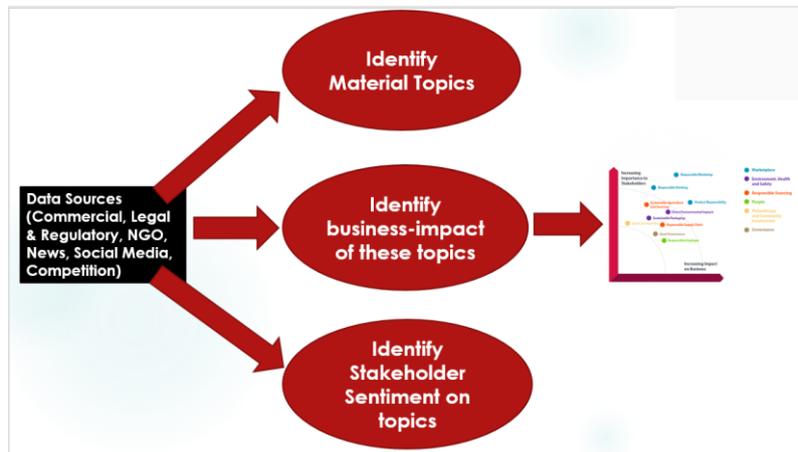
Fig. 23 – Context diagram of Tool

3.4   Competitor Datamaran by eRevalue

(London, New York, Valencia, San Francisco and Mumbai)
What
- semantic business intelligence tool
- aimed to resolve problems faced in monitoring and reporting risk
- Aims to provide fast, efficient and powerful business intelligence on large unstructured data sets from the corporate reporting landscape
- to help companies ensure resilience in a volatile market by having access to new insights through technology.

How
- based on the use of AI via Natural Language Processing (NLP) techniques
- process thousands of corporate reports identifying topics of interest and rank them based on the level of disclosure on each topic.
- capture sources of data including corporate filings, regulatory initiatives, social media, stakeholder opinion and news. The sources are curated to ensure the coverage is robust and quality is assured both geographically and across different business sectors. To the sources mentioned above NLP technology is applied focusing on the narrative to analyse a series of emerging issues. At this moment the user starts to interact with the data. Using a series of filters, a user creates a scenario based on their own business needs to build a view of the risk and opportunity across these emerging issues from a competitive, regulatory and reputational standpoint
- Datamaran also uses NLP to analyse stories from the news, utilizing named entities to identify companies that are being discussed and looking at trends in the topics that being frequently mentioned alongside these companies.

Details
- They currently perform their analytical services in English and Spanish. The SaaS platform is 100% cloud based, and they sell on a monthly and annual user license agreement. It provides visual analytics and sophisticated search 2010 till present, for £70 per month with their "Datamaran research" package. For £700 per month the "Datamaran Benchmark" package also provides their users with the following:
  1. Gap analysis
  2. Benchmark module
  3. Scenario building
  4. SDG mapping tool
  5. GRI mapping tool
- In addition, they have a free version which provides users with visual analytics and Sophisticated search from data 2010 till 2015. Their business intelligence tool captures data from over 50.000 Corporate Reports, 7000 companies and 3000 regulations.

The following insights were obtained:
The solution proposed by us should aim to first get existing customers on board, rather than acting as a competition to Datamaran or targeting the latter's customer base.
Our tool should focus extensively on step by step Materiality Assessment process/ consulting, rather than an overall Sustainability reporting tool.

Our tool should have options to work in more languages than Datamaran. For example, Datamaran operates in English and Spanish, and our tool would be operational in English, Dutch, German, French and Spanish to provide added value.

Initially, it would not be sold as a commercial product, but used as an asset to guide and streamline the sustainability consulting business. Basically, the aim would be to use it internally and reduce effort considerably and conduct internal validations. Rather than acquiring revenue, the approach should be to save costs.

Based on the above insights, a SWOT Analysis (Strengths, Weaknesses, Opportunities and Threats) were captured in the illustration.

STRENGTHS	WEAKNESSES
1. Step by step guidance for Materiality Assessment	1. Undertaken as a University Master's project on limited resources
2. Automation of majority of manual and time consuming activities in the process	2. Needs to be made robust for commercial implementation
3. Minimal human involvement, thus reduced subjectivity	
4. Scalable to accommodate any number of languages and large input volumes	
5. Ability to distinguish between various languages in input documents	
6. Created with R Studio allowing for easy extension of functionalities in future	
7. Easy visualizations for demonstrations	
OPPORTUNITIES	THREATS
1. Market as exclusively for step-by step Materiality Assessment tool	1. Since the project had academic motivations, some details might be publicly available to competitors
2.Use internally for cost saving by reducing effort for consulting	
3. Focus on existing Dutch sustainability reporting market	
4. Aim to capture upcoming markets in Germany and Belgium owing to existing clients therein	

Fig. 24 – SWOT Analysis

Appendix 4 – Requirements mapping to Features

After multiple iterations of the development and feedback, the first acceptable version of the tool was obtained. The mapping of the requirements to the features proposed is shown via screenshots in this appendix.

Pain-Point number	Pain-Point name	Explanation	Requirement (to resolve)	Feature	Location (in tool)	Screenshot number
PP1	Conduct Desk Research	Scanning through input documents, looking for links wih the Longlist	Document scanning to identify Phrase/term based on Longlist	Upload PDF files, Upload Longlist, Create PDF	Input panel, Tab - Word Clouds	1,2,3
PP2	Identify important topics			Text matching algorithm (runs in the backend)	Backend	NA
PP3	Populate Longlist (Term Document Matrix)	For each item in Longlist, indicate occurrence matches with respect to each input document	Create Term Document Matrix based on previous step	Create Term Document Matrix	Tab - Term Document Table	4
PP4	Gather suggestions	Obtain opinion from Company about the topics (the matches found in previous step)	Option to input suggestions from company for all the Longlist matches	Create Longlist word cloud	Tab - Word Clouds	1,2,3, 5
PP5	Shorten Longlist by results	Incorporate opinions from previous step in Longlist, thereby excluding some items and shortening the list	Filter out and shorten Longlist based on previous step	Option to select/ deselect Longlist items	Tab - Matrix	5
PP6	Analyse existing documentation	Scanning through second set of input documents	Document scanning to identify Phrase/term from shortened Longlist for second set of documents	Upload PDF files, Upload Shortlist	Input panel, Tab - Matrix	1, 6
PP7	Assign importance by stakeholder category	For each topic, assign weightage based on previous step	Option to assign weightage to the shortlist of topics obtained from previous steps	Manual input to shortened list of topics	Tab - Matrix	5
PP8	Translate results	Plot Materiality Matrix	Select number of topics to display, and plot them on graph according to weightage	Create Matrix (plot graph)	Tab - Matrix	5, 6

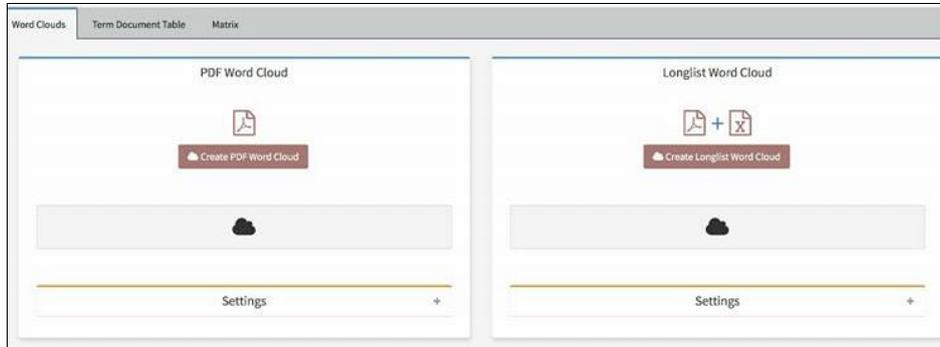**Table 10 –** Features mapped to screenshots in tool

Fig. 25 – Screenshot1 of Word Cloud tab

Fig. 26 – Screenshot2 of Word Cloud tab


Fig. 27 – Screenshot3 of Word Cloud tab

Categories	Longlist	Score	Boskalis_Ann_t_2016	CSR_Report	DEME_Activiti 2016
	Company competitiveness / competitive behavior / fair competition / anti-competitive behaviour	2	0	0	0
	Economic performance	1	0	0	0
	Long term viability of core business	0	0	0	0
	Market presence	0	0	0	0
	New markets / emerging economies	1	0	0	0,007
	Corporate governance / organizational governance	4	0,062	0,112	0
	Response to organizational change	0	0	0	0
	Board structure and independence	0	0	0	0
	Gender participation on governance bodies	0	0	0	0
	Executive compensation	0	0	0	0
	Business ethics	2	0	0	0
	Regulatory compliance	0	0	0	0
Business	Environmental compliance	0	0	0	0
	Regulatory and legal challenges	0	0	0	0
	Communications and engagement	0	0	0	0
	Shareholder engagement / Stakeholder engagement / Shareholder Expectation	4	0	0,038	0
	Innovation	6	0,062	0,05	0,169
	Innovation management process	0	0	0	0
	Sustainable innovations	0	0	0	0

Fig. 28 – Screenshot4 of Term Document Matrix tab

	Adjust scores			
	Longlist	External	Internal	Show
1	Policies / standards / codes of conduct	10.0	9.0	☑
2	Emissions / CO2 footprint	8	3.0	☑
3	Materials	8.0	7.0	☑
4	Human rights	5.4	2.0	☑
5	Biodiversity / eco persevation	5.4	8.0	☑
6	Innovation	4.6	9.0	☑
7	Emissions to air / GHG	4.6	6.0	☑
8	Due diligence / human rights assessments	4.6	3.0	☑
9	Corporate governance / organizational governance	3.1	5.0	☑
10	Shareholder engagement / Stakeholder engagement / Shareholder Expectation	3.1	8.0	☑
11	Occupational health / safety risks	3.1	1.0	☑
12	Grievance mechanism	2.3	4.0	☑
13	Company competitiveness / competitive behavior / fair competition / anti-competitive behaviour	1.5	2.0	☑
14	Business ethics	1.5	7.0	☑
15	Anti-corruption	1.5	9.0	☑
16	Fundamental principles / rights at work	1.5	0.0	☐
17	Child labour	1.5	0.0	☐
18	Pollution prevention	1.5	0.0	☐

Fig. 29 – Screenshot5 of Matrix tab
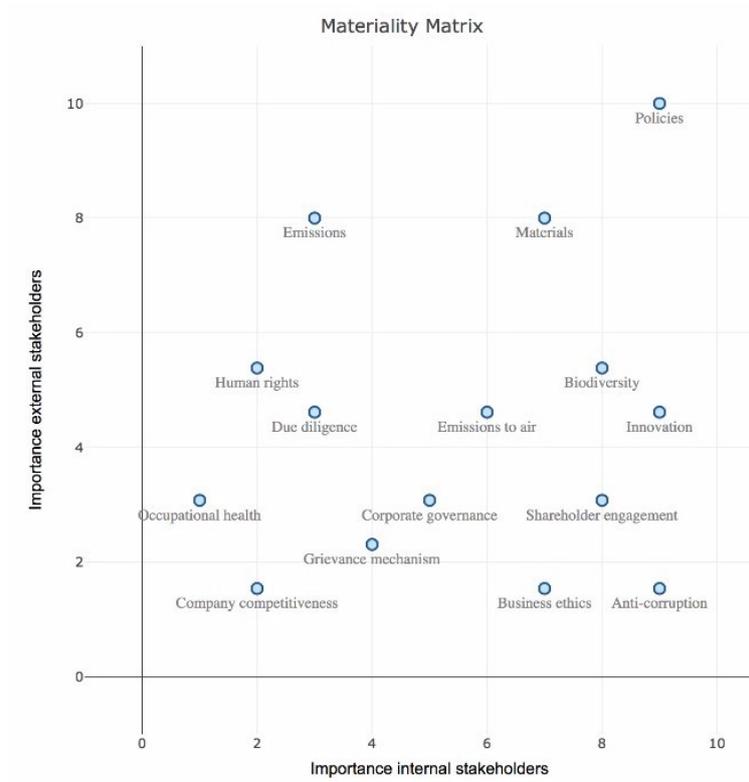
Fig. 30 – Screenshot6 of Matrix tab

Appendix 5 – Validation

Details of the validation process are captured in this section.

5.1 Perceived Parameters Validation

The Method Evaluation Model, used for Validation of the tool was adapted from Davis et al's (1989) study. The following variables were used:

- Perceived Ease of Use (PEOU) - This was measured using six items on the survey (Questions 1, 4, 5, 9, 11 and 14).
- Perceived Usefulness (PU) - This was measured using eight items on the post-task survey (Questions 2, 3, 6, 7, 8, 12, 13, 15).
- Intention to Use (ITU) - This was measured using two items on the post-task survey (Q10 and Q16).

The following questionnaire was circulated among prospective users of the tool:

Q1. I found the interaction with   for materiality assessment complex and difficult to follow

Q2. I believe that this tool would reduce the effort required to conduct large materiality assessment tasks

Q3. This tool would increase my performance when carrying out materiality assessment tasks

Q4. Overall, I found   difficult to use

Q5. This tool would make it easier for us to verify whether manual materiality assessments (e.g. assessments performed by interns) performed at our company are correct

Q6. I found   easy to learn

Q7. Overall, I found   to be useful

Q8. Using this tool would make it more difficult to succeed in conducting materiality assessment when there is a large quantity of input documents

Q9. I found it difficult to apply   to the samples of customer data

Q10. I would definitely not use this tool to conduct materiality assessment

Q11. I found the functionalities of   clear and easy to understand

Q12. Overall, I think this tool does not provide an effective solution to the problem conducting materiality assessment exercises

Q13. Using this tool would make it easier to communicate simulations of materiality assessment exercises to our customers

Q14. I am not confident that I am now competent to use this tool in practice

Q15. Overall, I think this tool is an improvement to the way we perform materiality assessment at our company

Q16. I intend to use this tool in preference to the manual way of working if I have to carry out materiality assessment in the future

The responses were measured on a 5-point Likert scale (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree).

Out of the 10 consultants the survey was sent to, we received 4 responses. Details of this are shown below.

Variable ID	Question number	Question	2017/10/27 3:07:35	2017/10/27 4:52:42	2017/10/30 10:31:00	2017/11/09 4:57:56	2017/12/05 1:26:14	2017/12/05 1:41:13	2017/12/05 1:49:28	2017/12/05 1:54:59
PEOU1	1	I found the interaction with the tool for materiality assessment complex and difficult to follow	2	2	2	2	1	1	2	2
PU1	2	I believe that this tool would reduce the effort required to conduct large materiality assessment tasks	5	3	5	5	5	5	4	4
PU4	3	This tool would increase my performance when carrying out materiality assessment tasks	5	4	4	5	5	5	4	5
PEOU2	4	Overall, I found the tool difficult to use	2	2	1	2	3	2	2	2
PU4	5	This tool would make it easier for us to verify whether manual materiality assessments (e.g. assessments performed by interns) performed at our company are correct	4	5	4	4	4	4	4	3
PEOU3	6	I found the tool easy to learn	4	3	4	5	3	4	4	4
PU4	7	Overall, I found the tool to be useful	5	4	5	5	5	5	5	5
PU5	8	Using this tool would make it more difficult to succeed in conducting materiality assessment when there is a large quantity of input documents	1	1	1	2	1	1	1	2
PEOU3	9	I found it difficult to apply the tool to the samples of client data	1	3	1	1	2	2	2	2
ITU1	10	I would definitely not use this tool to conduct materi	1	1	1	2	1	1	2	2
PEOU5	11	I found the functionalities of the tool clear and easy to understand	4	4	4	3	4	4	4	4
PU6	12	Overall, I think this tool does not provide an effective solution to the problem conducting materiality assessment exercises	1	3	1	2	1	1	1	2
PU5	13	Using this tool would make it easier to communicate simulations of materiality assessment exercises to our clients	5	4	4	4	4	4	4	3
PEOU6	14	I am not confident that I am now competent to use this tool in practice	2	4	4	3	2	2	2	2
PU4	15	Overall, I think this tool is an improvement to the way we perform materiality assessment at our company	5	4	4	5	4	5	5	4
ITU2	16	I intend to use this tool in preference to the manual way of working if I have to carry out materiality assessment in the future	5	3	5	5	5	4	4	4

Table 11 – Survey responses

The Polarity identification and calculation of mean scores done as part of the data preparation for analysis is shown below:

Variable Number	Question number	Variable	Variable inclination	Questions
PU6	12	Perceived Usefulness	Perceived Usefulness - negative	Overall, I think this tool does not provide an effective solution to the problem conducting materiality assessment exercises
PU5	8	Perceived Usefulness	Perceived Usefulness - negative	Using this tool would make it more difficult to succeed in conducting materiality assessment when there is a large quantity of input documents
PU5	13	Perceived Usefulness	Perceived Usefulness - positive	Using this tool would make it easier to communicate simulations of materiality assessment exercises to our clients
PU4	3	Perceived Usefulness	Perceived Usefulness - positive	This tool would increase my performance when carrying out materiality assessment tasks
PU4	5	Perceived Usefulness	Perceived Usefulness - positive	This tool would make it easier for us to verify whether manual materiality assessments (e.g. assessments performed by interns) performed at our company are correct
PU4	7	Perceived Usefulness	Perceived Usefulness - positive	Overall, I found the tool to be useful
PU4	15	Perceived Usefulness	Perceived Usefulness - positive	Overall, I think this tool is an improvement to the way we perform materiality assessment at our company
PU1	2	Perceived Usefulness	Perceived Usefulness - positive	I believe that this tool would reduce the effort required to conduct large materiality assessment tasks
PEOU6	14	Perceived Ease of Use	Perceived Ease of Use - negative	I am not confident that I am now competent to use this tool in practice
PEOU5	11	Perceived Ease of Use	Perceived Ease of Use - positive	I found the functionalities of the tool clear and easy to understand
PEOU3	6	Perceived Ease of Use	Perceived Ease of Use - positive	I found the tool easy to learn
PEOU3	9	Perceived Ease of Use	Perceived Ease of Use - negative	I found it difficult to apply the tool to the samples of client data
PEOU2	4	Perceived Ease of Use	Perceived Ease of Use - negative	Overall, I found the tool difficult to use
PEOU1	1	Perceived Ease of Use	Perceived Ease of Use - negative	I found the interaction with the tool for materiality assessment complex and difficult to follow
ITU2	16	Intention To Use	Intention To Use - positive	I intend to use this tool in preference to the manual way of working if I have to carry out materiality assessment in the future
ITU1	10	Intention To Use	Intention To Use - negative	I would definitely not use this tool to conduct materiality assessment

Table 12 – Survey response Analysis

Post the polarity identification, the scores for the negative questions were converted to positive, using the formula:

New score = (Lowest core + highest score) - original score

Negative score	Equivalent positive score
1	5
2	4
3	3
4	2
5	1

Table 13 – Negative to positive scores conversion

The table below shows the scores of all questions after conversion (the scores for the positive questions remain unchanged).

VariableID	Variable	Converted score respondent 1	Converted score respondent 2	Converted score respondent 3	Converted score respondent 4	Converted score respondent 5	Converted score respondent 6	Converted score respondent 7	Converted score respondent 8
PEOU1	PEOU	4	4	4	4	5	5	4	4
PU1	PU	5	3	5	5	5	5	4	4
PU4	PU	5	4	4	5	5	5	4	5
PEOU2	PEOU	4	4	5	4	3	4	4	4
PU4	PU	4	5	4	4	4	4	4	3
PEOU3	PEOU	4	3	4	5	3	4	4	4
PU4	PU	5	4	5	5	5	5	5	5
PU5	PU	5	5	5	4	5	5	5	4
PEOU3	PEOU	5	3	5	5	4	4	4	4
ITU1	ITU	5	5	5	4	5	5	4	4
PEOU5	PEOU	4	4	4	3	4	4	4	4
PU6	PU	5	3	5	4	5	5	5	4
PU5	PU	5	4	4	4	4	4	4	3
PEOU6	PEOU	4	2	2	3	4	4	4	4
PU4	PU	5	4	4	5	4	5	5	4
ITU2	ITU	5	3	5	5	5	4	4	4

Table 14 – Final scores after conversion

The results of the normalization for each variable are shown in the tables 15, 16 and 17 below.

VariableID	Variable	Respondent ID							
		1	2	3	4	5	6	7	8
PEOU1	PEOU	4	4	4	4	5	5	4	4
PEOU2	PEOU	4	4	5	4	3	4	4	4
PEOU3	PEOU	4	3	4	5	3	4	4	4
PEOU3	PEOU	5	3	5	5	4	4	4	4
PEOU5	PEOU	4	4	4	3	4	4	4	4
PEOU6	PEOU	4	2	2	3	4	4	4	4
	Sum	25	20	24	24	23	25	24	24
	Normalised score in %	79%	58%	75%	75%	71%	79%	75%	75%

Table 15 – Normalized scores, Perceived Ease of Use

VariableID	Variable	Respondent ID							
		1	2	3	4	5	6	7	8
PU1	PU	5	3	5	5	5	5	4	4
PU4	PU	5	4	4	5	5	5	4	5
PU4	PU	4	5	4	4	4	4	4	3
PU4	PU	5	4	5	5	5	5	5	5
PU5	PU	5	5	5	4	5	5	5	4
PU6	PU	5	3	5	4	5	5	5	4
PU5	PU	5	4	4	4	4	4	4	3
PU4	PU	5	4	4	5	4	5	5	4
	Sum	39	32	36	36	37	38	36	32
	Normalised score in %	97%	75%	88%	88%	91%	94%	88%	75%

Table 16 – Normalized scores, Perceived Usefulness

VariableID	Variable	Respondent ID							
		1	2	3	4	5	6	7	8
ITU1	ITU	5	5	5	4	5	5	4	4
ITU2	ITU	5	3	5	5	5	4	4	4
	Sum	10	8	10	9	10	9	8	8
	Normalised score in %	100%	75%	100%	88%	100%	88%	75%	75%

Table 17 – Normalized scores, Intention to Use

Table 18 below shows the summary of the normalized scores across all variables.

Normalised score in %	PEOU	PU	ITU
Respondent 1	79%	97%	100%
Respondent 2	58%	75%	75%
Respondent 3	75%	88%	100%
Respondent 4	75%	88%	88%
Respondent 5	71%	91%	100%
Respondent 6	79%	94%	88%
Respondent 7	75%	88%	75%
Respondent 8	75%	75%	75%
Average	73%	87%	88%

Table 18 – Normalized scores, All variables

Appendix 6 – References for future work

This section contains starting points for future work on this tool. The template embedded here has a potential list of sources that may be used to capture materiality data. The second sheet has a suggested scoring mechanism to rate each score based on factors like coverage, reliability, and availability.

Sources.xlsx