IFC Workshop on "Addressing climate change data needs: the global debate and central banks' contribution"

6-7 May 2024

Combining AI and domain expertise to assess corporate climate transition disclosures[1]

J Bingler,
University of Oxford / Council on Economic Policies, Zurich

M Leippold,
University of Zurich / Swiss Finance Institute

J Ni,
University of Zurich / ETH Zurich

T Schimanski and C Senni,
University of Zurich

[1] This contribution was prepared for the workshop. The views expressed are those of the authors and do not necessarily reflect the views of the Central Bank of the Republic of Türkiye, the BIS, the IFC or the other central banks and institutions represented at the event.

Combining AI and Domain Expertise to Assess Corporate Climate Transition Disclosures

Chiara Colesanti Senni,[1*] **Tobias Schimanski,**[1†] **Julia Bingler,**[2, 3] **Jingwei Ni,**[1,5] **Markus Leippold**[1, 4]

[1]Department of Finance, University of Zürich, Zürich, Switzerland
[2]Smith School of Enterprise and the Environment, University of Oxford, Oxford, United Kingdom
[3]Council on Economic Policies, Zürich, Switzerland
[4]Swiss Finance Institute (SFI), Zürich, Switzerland
[5]Center of Economic Research, ETH Zürich, Zürich, Switzerland

Abstract

Companies need sound planning to reduce their emissions and deal with the transition to a more sustainable economy. The disclosure of such plans is key for effective capital allocation and risk management. Transition and sustainability disclosures are a compass for market participants to guide their actions and strategies toward the net-zero target. If companies plan their transition appropriately, the negative implications of physical and transition risks for micro- and macro-financial stability can be reduced. Many frameworks have been suggested to assess transition plans' ambition, credibility, and feasibility. However, the lack of one clear reference framework paves the way for inconsistencies in transition plans and the risk of greenwashing. We propose a set of 64 common ground indicators from 28 different transition plan disclosure frameworks to comprehensively assess transition plans and develop a novel natural language processing (NLP)–based tool to automate the assessment of companies' disclosures. This can help investors and financial supervisors assess transition risks while supporting companies' disclosure efforts. Applying the tool to 143 reports from the carbon-intensive CA100+ companies, we find that companies tend to disclose more indicators related to target setting (talk) but fewer indicators related to the concrete implementation of strategies (walk). Our results demonstrate that machine learning can be used to generate a positive impact on the transition towards a more sustainable economy by identifying the elements of transition plans that require further scrutiny and/or effort. Our work will be a starting point for further leveraging new technologies in sustainable finance. For example, the assessment of the plans could be used by financial regulators in their supervisory practices or to investigate whether the risk of greenwashing is reflected in stock returns.

Keywords: Climate disclosure, RAG system, transition strategies, human evaluation, CA100+.

[*]Corresponding Author (chiara.colesantisenni@df.uzh.ch)
[†]Corresponding Author (tobias.schimanski@df.uzh.ch)
First version: November 2023
This version: May 13, 2024

Introduction

As the impacts of climate change become increasingly severe, the urgency for the global community to transition to a net-zero economy has never been more critical. This transition is essential for mitigating the negative effects of climate change and ensuring economic stability and sustainability in the long term. Corporations, as significant contributors to greenhouse gas emissions, are central to this transformation. They are tasked with adjusting their operations and strategies to align with climate goals. This requires comprehensive planning, implementation, and transparency. Furthermore, this transformation requires a large amount of financial resources. Financial institutions are pivotal in directing capital towards sustainable activities and enabling the transition of corporate business models and technologies. Yet, they need information on comprehensive corporate transition strategies to assess the future risks and opportunities associated with their investments.

Ambitious, credible, and feasible transition strategies are relevant not only for corporates- which use them to navigate the shift towards sustainable practices and ensure that they remain competitive and compliant with an evolving regulatory landscape- but also for a wide range of stakeholders. For investors, these plans gauge a company's long-term viability and alignment with increasingly stringent environmental standards, such as those set out in the Paris Agreement. Regulators rely on these plans to ensure that companies make genuine efforts to reduce their carbon footprints, which is critical to meeting national and international climate targets. Hence, the evaluation of transition strategies is key, as shown by recent contributions in this direction [1,2,3,4,5].

Despite the critical importance of corporate transition strategies, several challenges can undermine their effectiveness in supporting the achievement of the net-zero target. Reporting and credibility issues whereby companies misleadingly portray their climate and environmental efforts (greenwashing) are at the forefront. This can undermine trust and lead to a misallocation of resources in the economy, putting the climate and environmental targets at risk and coming along with negative micro- and macro- financial stability implications [6,7,8,9,10,11]. The lack of standardization in companies' disclosures reduces stakeholders' ability to compare and assess the ambition, credibility, and feasibility of the transition strategies of different companies. This poses a substantial barrier to effectively evaluating corporate contributions to climate goals and strategies to steer their business to the future. At the same time, the ongoing pressure to show progress in reducing climate impact leads companies and public institutions to release information that often results in vast amounts of unstructured data about transitioning toward net zero [12].

This paper addresses these challenges by providing a practical framework that assesses corporate climate transition disclosures. Our approach uses a natural language processing (NLP) tool to automate and enhance the analysis of sustainability disclosures and identify potential risks of greenwashing. We build on previous work that uses NLP in the field of sustainable finance and corporate climate risk analysis e.g., [13,9,14,15,16,17] and extend the literature by providing a standardized, detailed, and expert-driven scaleable assessment process. Thus, we respond to the call for more fine-granular AI assessments in sustainability [18].

Our contribution is threefold. First, we identify the common ground of comprehensive indicators to assess transition strategies. To achieve this goal, we review 28 different transition strategy disclosure frameworks to identify common criteria for assessing corporate transition strategies and elicit experts' opinions to develop a unified framework encompassing 64 indicators. These indicators cover different key areas for a comprehensive, successful net-zero transition. In this context, the lack of disclosure of the information required by individual indicators represents a potential risk associated with the company, as it signals that it might not be well-prepared for the transition in that specific dimension.

Second, we build and validate an NLP tool for the automated analysis of transition plans based on the 64 common ground indicators. We assess the performance of the NLP tool in a pilot evaluation involving users from 26 different institutions to gain insights into practitioners' perceptions of the tool, its trustworthiness, and practical usage. The pilot involved representatives from both the private and public sectors and profit and non-profit organizations. We find that users are very satisfied with the tool's performance overall, but there is use-case specific disagreement regarding the priorities for improvement.[1]

Third, we investigate the disclosures of the Climate Action 100+[2] companies to identify potential inconsistencies in their transition strategies. We focus on these companies, as they are critical for the transition to a low-carbon economy at the global level. Our findings show that companies tend to disclose more indicators related to target setting (talk) but fewer indicators related to the concrete implementation of strategies (walk).

Our research contributes to the fields of corporate climate risk analysis, sustainable finance, and

[1] You can try out the tool on our GitHub repository. In there, you will find the code and data of this project as well as further helpful resources.

[2] Climate Action 100+ is an investor-led initiative that aims to incentivize the world's largest corporate greenhouse gas emitters to take climate action.

corporate governance by enhancing the ability to monitor and assess the ambition, credibility, and feasibility of corporate transition plans. It offers new insights and solutions that can be adopted by policymakers, financial supervisors, corporations, and financial institutions to support a more resilient and sustainable economic and financial system.

1 Assessment indicators

We initially define a common ground of assessment indicators along which companies' disclosure related to transition strategies can be analyzed. To identify these assessment indicators, we proceed in three steps. First, we review existing transition plan disclosure frameworks and extract the most common criteria for sound transition strategies. Second, we combine quantitative and qualitative evaluations to create a structured list of the most important indicators. Third, we evaluate the importance of each indicator by collecting feedback from a group of stakeholders involved in the transition plan analysis and define the final list of indicators based on their comments.

In the first step, we review 28 existing transition plan frameworks and identify commonly suggested assessment indicators. These indicators are criteria to evaluate the ambition, credibility, and feasibility of transition plans.[3] In this study, we review frameworks that were published over the period 2021-2023 by different stakeholders and initiatives (see Table S.1 for an overview).

In the second step, we use quantitative and qualitative criteria to assess the importance of the identified indicators for ambitious, comprehensive, and credible corporate transition strategies. We first compute the frequency of indicators' appearance in the reviewed frameworks for the quantitative evaluation. In particular, we count the number of times one indicator is considered in the assessed frameworks and assign one value for each appearance. In some cases, we assign a value of 0.5 if an indicator is only partially covered by the respective framework. For example, an indicator could be a recommendation amongst others and not a core required element. This quantitative assessment helps to obtain an initial weight of the importance of an indicator. Moreover, we undertake a qualitative assessment by discussing the quantitative weights for each indicator, combining similar indicators, and verifying their suitability for assessing transition disclosures. In addition, we classify our indicators into "walk" (W) and "talk" (T). The distinction is made based on whether a specific indicator relates to future targets and/or general transition monitoring and management approaches (T) or to specific and already verifiable transition activities (W).

In the third step, we share our list of indicators with more than 50 selected experts who formed our advisory board. The advisory board included financial industry representatives, central bankers, and financial supervisors.[4] The experts were asked to provide comments, suggest amendments, and refine the indicators. We had both group meetings with the whole advisory board and individual meetings to understand the main criticalities and the importance of different indicators to different stakeholders. As a result, we obtain a common ground list of 64 indicators along which we assess transition-related disclosures.

Section S.3 lists all the indicators selected, covering the broad categories "Target", "Governance", "Strategy", and "Tracking" (as used in e.g. TCFD). In addition, we show each individual indicator's classification as primarily a walk (W) or talk (T) indicator.

With our framework, it is possible to analyze whether companies disclose information related to specific indicators or not. For each not-disclosed indicator, our approach signals that the company's transition strategy might be entirely absent or inconsistent in that specific dimension (for example, not ambitious, credible, or feasible). As such, further investigation by stakeholders and disclosure efforts by the respective company are required.

An important advantage of the proposed framework is its flexibility: While we suggest a list of indicators deemed appropriate to assess transition strategies, users can modify and extend this depending on their needs. For instance, financial institutions lending to specific sectors might want to add more details about risks that are specifically important for their assessment. Similarly, users interested in broader nature-related risks might want to include additional questions about supply chains and companies' locations.

2 NLP Model Development

We develop a natural language processing (NLP) tool to automate the analysis of company disclosures along with our indicators. Furthermore, we validate the tool in a pilot study with 26 institutions.

NLP Tool

To translate our framework into an automated analysis tool, we rely on Large Language Models (LLMs). LLMs have shown vast capabilities in reasoning, understanding, and generation of text[19,20,21,22]. However, LLMs also face two major

[3] The literature also refers to the different dimensions of consistency as internal (ambition and feasibility) and external (feasibility).

[4] The advisory board was established to provide feedback throughout the entire project. For more details on the advisory board, see Section S.1.

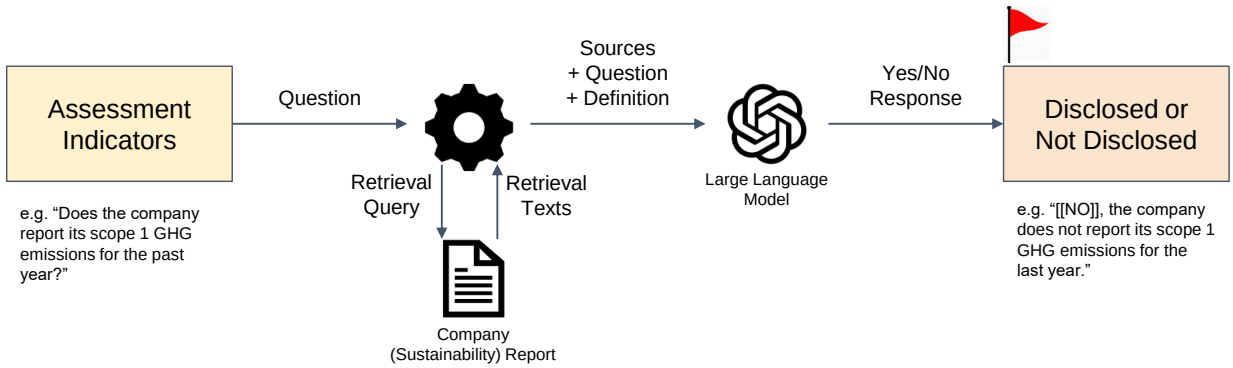


Figure 1: Retrieval Augmented Generation (RAG) pipeline for analyzing the assessment indicators.

challenges that arise from relying on internal knowledge: (1) LLMs may produce hallucinated output, that is, text that is not factual[23], and (2) the embedded knowledge base of LLMs is intransparent and truncated at a certain point in time. Against these shortcomings, practitioners and researchers rely on a technique called Retrieval Augmented Generation (RAG)[24]. RAG systems aim to include external knowledge in the prompts provided to the LLM and force the model to rely only on this information when answering a given question. Thus, RAG systems make use of the strong capabilities of LLMs to summarize and reason over the provided content and try to minimize the dependence on the internal knowledge of the LLM.

Figure 1 displays the RAG system used in this project. The starting point is given by the assessment indicators in the form of questions. Following the suggestions from our advisory board, we aim for an easy-to-use and understand approach, as well as a high-level yes/no answer for the assessment indicators. This entails a binary decision by the tool on whether the information required by the indicator is disclosed or not. For example, one indicator considered is related to the disclosure of scope 1 GHG emissions in the past year. Hence, the corresponding question reads: "Does the company report its scope 1 GHG emissions for the past year?". If the NLP tool finds information related to past year emissions in a company's disclosures, it will return a YES followed by an explanation about the decision made.

From a technical perspective, each question is embedded in a numerical representation. Accordingly, the disclosures of the investigated company are split into chunks, and every chunk is embedded in a numerical representation. Chunks with similar semantics obtain similar numerical representations. As a result, the numerical representation of the question can be used to find semantically similar, that is, question-relevant texts. These retrieved texts are included in a structured prompt format and provided to the LLM, together with specific guidelines that help pin down the exact information on which to focus when answering the question. We defined the guidelines in an iterative, expert-based process by assessing the model answers for each indicator based on the simple question and adding further information as needed until the answers were comprehensive and in line with the experts' knowledge and assessments (see Appendix for more details S.5). Furthermore, the prompt includes formal instruction and, importantly, the command that the LLM should only rely on the given sources and not on internal knowledge. This is important to ensure that the model only assesses the specific corporate document that is to be analyzed for the user and does not draw on information about the company that was available elsewhere when the model was initially trained.

Finally, the tool produces a structured answer indicating whether information as requested by the indicator under investigation, formulated as a question, is available. The output is a yes or no answer, followed by an explanation of the decision and the source references based on the answer (see Section 5 for technical details). These explanations can enable a more holistic and, importantly, transparent understanding of the evaluation made. They allow the user to understand the reasoning behind the choice and can provide starting points for more detailed investigations. The sources allow the user to cross-check whether the relevant information has been extracted by the model and whether it has been correctly summarized (see Section S.6 for more details).

Human Evaluation

We evaluate our tool in a pilot study. We choose a tool evaluation design that helps us obtain quantitative and qualitative feedback from domain experts. This adds to prior research, which has mainly addressed the quantitative evaluation of RAG systems in artificial or theoretical setups[25,26,27,28]. Only a few analyses have considered expert-based evaluations[29,30]. However, none of these prior pa-

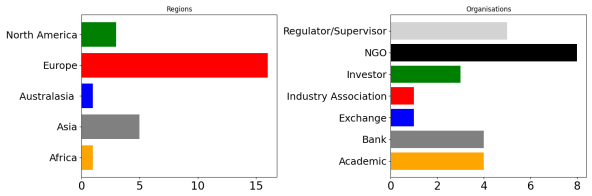


Figure 2: Overview of regions and organization types of the 26 participating organizations to the study.

pers attempted a comprehensive evaluation of quality dimensions, such as correct sourcing and answering, as well as the potential usage of such tools for stakeholders.

Our evaluation incorporates feedback from domain experts from 26 organizations. The experts include financial regulators and supervisors, investors, exchanges, NGOs and industry associations representatives, bankers, and academics. The participants are predominantly based in Europe, followed by participants from Asia, North America, Australasia, and Africa.[5] See Figure 2 for more details. One organization participated with three persons. Thus, we have 28 participants in total.

The domain experts assess the tool along the following qualitative dimensions:

System Quality, which aims at understanding if relevant information is retrieved, answers are accurate and faithful, and reasoning capabilities are solid.

Trustworthiness, which yields insights into users' perceived trustworthiness and expertise of the model.

Usage, which aims at understanding the usefulness of the tool for the respective stakeholder, as well as future possible use cases.

In the evaluation, we differentiate between feedback on the tool's assessment of individual indicators, that is, the responses and sources for specific yes / no questions (Q1-Q9), and overall feedback about the tool in general (Q10-Q15). The participants provided an answer based on a given set of options (multiple choice) and could provide additional explanations (free text). For the detailed questions and setup, see Supplementary Material S.7.

We ask the pilot participants to submit at least three reports of interest for evaluation. This ensures that the experts have prior experience in analyzing the underlying companies. We obtained a total of 93 reports from our participants (see Figure S.7 for details on the reports), which we analyzed with our NLP tool. Then, we identify a pool of 12 indicators, which we consider representative and important to get a solid understanding of the tools' performance in assessing the available information. For each of the three reports the participants submitted, we randomly assigned two of the 12 indicators to each participant to assess. This ensures that participants assess at least two answers in detail and that we obtain a reasonably large cross-section of assessments across participants for the assessed indicators. Participants are free to assess more responses in addition to the assigned ones. Finally, each participant is requested to provide a general tool assessment. As a result of this process, we obtain 28 tool assessments and 396 assessment indicator evaluations.

The results of the human evaluation offer several insights.[6] Regarding **system quality**, participants display a very high level of satisfaction. For example, 81% of the participants found that the model summarizes relevant content correctly for the indicator assessments without making up information not contained in the report, while 12% found that this is partially the case, and only 6% were not satisfied with the result in this regard (Q4). Furthermore, the majority of participants perceive that the tool captures the most relevant content for the requested indicator information (Q3) and cites it correctly without making sources up (Q6-Q9). Amongst those unsatisfied with the model output for the indicator assessed, feedback included that the model may be too judgmental or occasionally missed out on important sources for the relevant information in the report on the specific indicator. When asked what the major area of improvement for the model could be to support the usage, the responses were polarized. In particular, one part of the experts suggested a more detailed assessment, while another part highlighted that a broader overview would be desirable (Q14). The results related to the system quality dimension align with prior research that outlines that the models can achieve very satisfactory quantitative results in faithfulness[28]. However, our analysis also reveals that for specific users, adaptation to individual needs is critical. This highlights that it is very difficult to find a one-fits-all solution. Rather, users should be able to adjust the tool to their specific needs.

Regarding **trusthworthiness**, the user's first impression of the individual answers is largely positive (Q2), and the sources provided in an answer largely support the trust in the model (Q5). This is also mirrored in the general tool assessment, where only three respondents claimed that they did not

[5] Although we have tried to achieve an acceptable level of representations in our sample, we had to rely on the self-selection of participants, which can influence the results of our pilot.

[6] We are aware that these answers might exhibit some bias. Although we ensured the anonymity of the reports submitted and the assessments provided, the responses of the self-selected participants might have been influenced by the setup of the pilot.

trust the information when using the model (Q15). Only two participants find the tool unsatisfactory (Q10). Generally, the answers of the tool were perceived as comparable with an expert with 1-2 or more years of experience in assessing companies' disclosures (Q11). However, it also becomes apparent that a large share of the participants is not sure about whether they would fully trust the information provided by the tool (11 of 28 respondents "partially" trust the model (Q15)). The optional explanations help us understand where these concerns stem from. Some participants outlined the lack of ability of the tool to handle sector-specific issues based on the current set of indicators. Others highlighted the need to compare the answers produced by the tool with third-party data.

Finally, the **usage** dimension sheds light on the ease of understanding the indicators themselves, use cases for applying the tool in practice, and improvement potentials. With respect to understanding the information requested by the indicators, the dominant feedback from the users is positive. Most of the participants would know which information they would look for if they had to answer the question themselves (Q1). Furthermore, stakeholders would utilize the tool for a wide set of use cases, ranging from corporate risk and opportunity assessment to high- and deep-level understanding of transition plans (Q13). As outlined in the System Quality dimension, enhancements reflect very individual needs (see Q14). A common qualitative feedback is that the tool can be valuable for quick assessment and understanding, but future improvements could include deeper and more actionable insights or sector-specific adjustments of the tool (Q12, Q16).

Overall, human evaluation provides valuable insight into understanding that the NLP tool represents a valuable asset for end users. However, the introduction of more user-specific adaptions of the tool could foster both trust and usage.

3 CA100+ Companies Use Case

To illustrate a use case of our tool, we apply it to analyze corporate sustainability reports of the CA100+ companies and assess the transition-related information contained in these reports along the 64 assessment indicators. CA100+ is an investor-led initiative to track the most-emitting companies in the world. Our sample covers 143 companies and the corresponding corporate sustainability reports for the fiscal year 2022. The distribution of the companies across the different sectors is shown in Figure 3. For all the reports in our sample, we assess how many indicators are disclosed in the report of the company under consideration (i.e., the tool assigned a "yes" answer to the prompt question of whether the information requested in the indicator is available in the report). Thus, for this use case, we focus on the quantitative part of the output. The qualitative answers to each indicator may provide additional information for users (see Section S.6 for a better understanding).

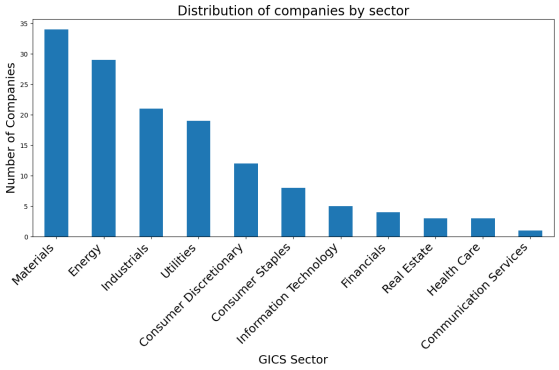


Figure 3: Distribution of companies by sector.

We find that the average count of indicators disclosed per report is 23 out of the 64 required indicators, while the best-performing report reaches a value of 43. The distribution of the average share of disclosed indicators by the company shows that a value of almost 40% is achieved by more than 35% of the companies (see Figure 4). However, it is important to understand the type of indicators for which information is available. Similarly, there might be sectoral patterns in the disclosure. Hence, we also rank the indicators according to their frequency of disclosure in the reports analyzed, and we look at sectoral disclosures.

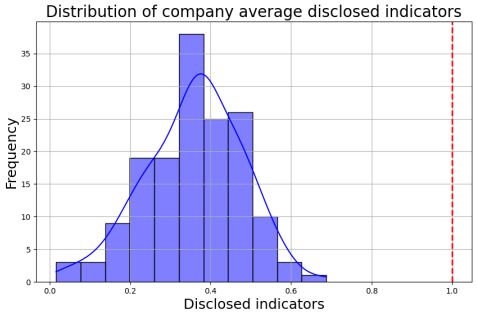


Figure 4: Distribution of the share of disclosed indicators by company.

Most and least disclosed indicators

In our analysis of corporate disclosures, we observe a clear divergence in disclosure depending on the types of indicators considered. This highlights areas where companies excel and fall short in their disclosures. The ranking of the disclosed indicators shows a pattern for the 10% most disclosed and the 10% least disclosed indicators (see Tables 1 and 2).

Identifier	Question	
9	Does the company report its GHG emission reduction interim targets for achieving the overall goal?	T
13	Does the company explain its governance structure for managing the climate transition?	T
17	Does the company report how its board oversees the climate transition plan implementation?	T
23	Does the company report quantitative or quantifiable subtargets in line with their climate targets and their climate key performance indicators?	T
26	Does the company report a renewable energy strategy and activities, covering renewable energy build out, procurement and consumption?	W
46	Does the company report its scope 1 GHG emissions for the past year?	W
50	Does the company report its annual progress of reducing GHG emissions to achieve its emission reduction or net zero targets?	W

Table 1: 10% most disclosed indicators

Identifier	Question	
12	If carbon credits and offsets are reported to be used by the company, does the company state explicitly that carbon credits and offsets will be only used when the company can ensure that the emission reduction or emission avoidance is sustained permanently?	T
19	Does the company provide a higher share of remuneration and bonuses that are linked to the successful implementation of the climate transition plan interim targets compared to the general part of variable compensation for executives and managers?	W
22	Does the company provide comprehensive evidence that it fully and completely integrates its climate strategy into its business strategy, product development, operations, financial and human resources, asset management and asset decommissioning?	W
36	Does the company report serious consequences and escalation strategies if net zero engagement is ineffective at upstream, downstream, policy maker and industry association level?	T
37	Does the company state explicitly that it stopped or will immediately stop any support or activities in new additional fossil fuel exploration and extension of fossil fuel supply?	W
38	Does the company report a strategy and activities for the decommissioning and canceling of planned or existing fossil fuel exploration and supply infrastructure?	W

Table 2: 10% least disclosed indicators

We first turn our attention to the top 10% disclosed indicators, that is, those indicators that have been disclosed the most by companies. These are related to GHG emission reduction interim targets, the structure of governance for environmental initiatives, oversight by corporate boards, and the specification of quantitative sub-targets. Moreover, companies tend to disclose assessment indicators of operational adjustments toward sustainability, including strategies related to adopting renewable energy and reporting Scope 1 GHG emissions from the past year.

In contrast, the disclosure of indicators that cover more comprehensive and ambitious implementation of transition strategies across all operations and activities is less frequent. These least disclosed indicators are related to the responsible use of carbon credits and offsets, alignment of executive remuneration with climate goals, and holistic integration of climate strategies across all business operations. The lack of disclosure suggests either a lack of fully developed strategies or a reluctance to reveal comprehensive details.

In addition, assessment indicators focusing on policy engagement transparency and ending the use of fossil fuels are less disclosed. These include examining the company's strategies for engagement in net-zero initiatives, policies on quitting the support for additional fossil fuel exploration activities, and plans to decommission fossil fuel infrastructure. The lack of disclosure in these areas could indicate a significant gap between companies' stated targets and the actual practices.

The divergence between the most and least disclosed indicators is also evident in terms of the classification of indicators into "walk" and "talk". Most of the least disclosed indicators are "walk" related, while the majority of the most disclosed indicators are "talk" related.

Sectoral analysis

Other than by the specific indicators, heterogeneity in the disclosures might also stem from sectoral characteristics. Hence, we investigate whether companies in some sectors disclose, on average, more information than others, as requested by our indicators.

Interestingly, we find that companies in sectors with the highest direct carbon emissions also have a higher amount of disclosed indicators. This could either mean that these companies are taking the transition strategy more seriously, that they are under higher pressure from investors and stakeholders to disclose their strategies, or that they are potential greenwashers.[7]

To account for potential differences in most and least disclosed indicators within sectors, we also look at the top and bottom 10% disclosed indicators in the Utilities, Energy, and Consumer Staples sectors. For these cases, we find similar results to those in the overall sample. For instance, the structure of governance for environmental initiatives and the oversight by the corporate board are always the most disclosed indicators, while the alignment of executive remuneration with climate goals and the decommissioning of fossil fuel infrastructures are the least disclosed.

The observed patterns in corporate disclosures suggest that companies may strategically prioritize reporting on areas that highlight their achievements and on communicating overall targets and less on the planned activities and changes in the business model and operations required to actually meet their future targets. This selective approach could be a form of greenwashing, as companies may highlight future promises together with easily achievable or lower impact aspects of their activities instead of stating faithfully their plans about core areas such as quitting the use of fossil fuels or their actual progress in reducing emissions across all scopes.

Again, our results hint at the fact that companies tend to disclose more indicators related to target setting (talk) but fewer indicators related to the concrete implementation of strategies (walk). This is in line with the findings of previous research and highlights the importance of calls for companies and the financial sectors to walk the talk[31,32].

4 Conclusion

Natural language processing offers unique opportunities for sustainable finance and climate-related financial risk management, including the assessment of corporate transition strategies with the aim to identify inconsistencies and potential areas of greenwashing. To the best of our knowledge, this is the first paper to show how a structured expert-centric NLP application process can help to identify a common analysis ground and assess the growing amounts of unstructured corporate climate transition information disclosed by companies.

More specifically, we first identified a common ground of assessment indicators to detect inconsistencies in transition strategies at the corporate level. We further introduce and evaluate an NLP tool to assess companies' disclosure and identify critical dimensions in the planning of the transition. Finally, we provided insights into climate transition disclosures of the world's most-emitting companies.

[7] The higher disclosure in some sectors might also be driven by the fact that some of the indicators are not relevant for all sectors (e.g., those related to the decommissioning of fossil fuels). However, the large majority of the selected indicators are sector agnostic (only three indicators are strictly related to fossil fuels and hence might not apply to all companies in the sample).

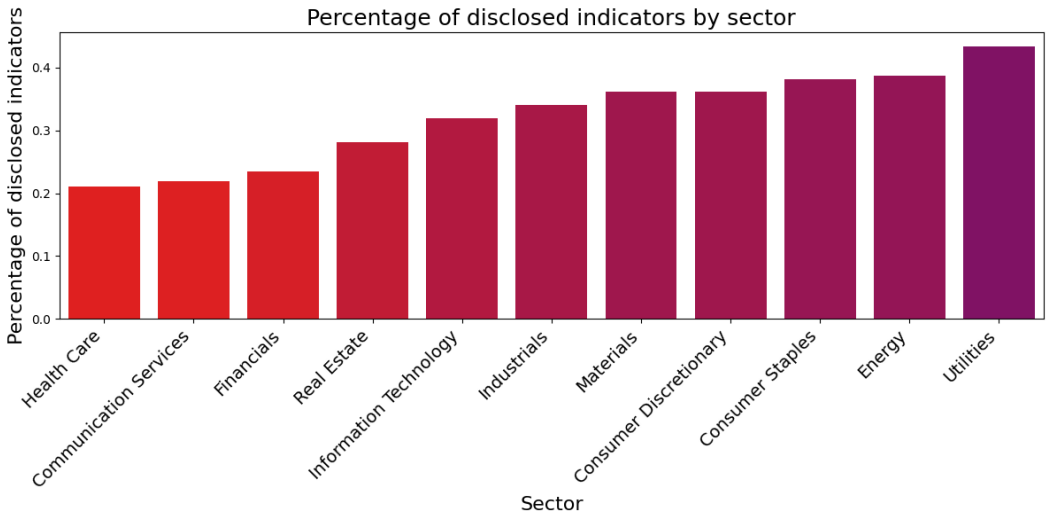


Figure 5: Average share of disclosed indicators by sector.

The examination of corporate disclosures along the climate transition strategy indicators with our NLP tool reveals that companies tend to disclose information related to target setting (talk), but less indicators related to the actual implementation of strategies (walk).

Our NLP tool aims to foster transparency in sustainability reporting and support the analysis of corporate climate transition strategies for research and practice. By obtaining detailed and verifiable disclosure assessments, stakeholders – including investors, regulators, and civil society – can more effectively evaluate the ambition, credibility, and feasibility of corporate climate transition strategies and advocate for more effective climate action in a targeted manner. Furthermore, an automated, consistent, and replicable method can assist company reporting and thereby enhance information availability and transparency for financial markets.

5 Limitations

Improving human evaluation

While the human evaluation represents a critical and valuable approach to understanding the expert-centric dimensions of NLP tools, our evaluation is not exhaustive and subjective in nature. More granular options for the answers would improve the elicitation of expert opinions. Similarly, more structured methods to elicit information could be introduced, such as Directed Acyclical Graphs. Moreover, we relied on participants' self-selection for the pilot study. Although we ensured anonymity in the responses, their choices and comments might have been influenced by having personally been in touch with the authors of this paper. The same limitation applies to the members of the advisory board.

Greenwashing

The tool cannot be used as a direct indicator of greenwashing. For instance, there might be specific company characteristics or regulatory contexts that explain the outcome. To more accurately capture greenwashing, the information disclosed should be compared with the information available in third-party datasets. This external information would allow users to properly verify whether what the company states aligns with its estimated transition and physical risk.[8]

Sector and country-specific assessments

In the current version, our tool compares the levels of ambition and the feasibility of implementing the transition strategy with global ambition goals and the internal consistency of the disclosed activities. However, for a more granular, appropriate assessment of ambition, sector-region-specific benchmarks would need to be applied. In addition, as highlighted by various researchers and think tanks currently working on credibility and feasibility assessments of transition plans,[9] Country-specific policies have a strong influence on the external feasibility drivers for corporate transition

[8] Clearly, third-party datasets are not a perfect solution as they are often based on companies' disclosures and can contain mistakes. However, the comparison would represent an improvement relative to an assessment based on our tool only.

[9] See for example policy briefs from Shrimali, et al. 2024, and an increasing amount of sector-country specific transition plans assessments like the sector-specific assessments aligned with France's NDC (ADEME, 2024) or the guidelines for assessing transition plans by company sectors from the Transition Plan Taskforce (TPT, 2024).

strategies. As a next step, it would be desirable to ground the tool in more granular country and sector-region-specific information for a more fine-tuned analysis.

Data and Model

Our model is naturally confined by the language and the level of detail provided in the disclosures analyzed. For the present analysis, this entails two limitations. First, our study solely focuses on sustainability reports. Thus, we may miss relevant information that a company presents elsewhere. However, the tool is applicable to any textual source, and different sources could be combined to analyze one company. Second, as the human evaluation shows, the tool is very satisfactory but not perfect. This implies that our model should be used as a complement and not as a substitute for manual analysis and verification processes. This is also in line with our general expert-centric approach, where we aim to identify the most useful human-machine collaboration methods. In this regard, our tool can meaningfully support human efforts by reducing the manual workload and help focus on specific areas of interest where companies disclose least in the assessment.

Current set of indicators

Depending on the scope, users might want to include different indicators in the analysis. This might be particularly important if users want to focus on specific sectors for which only some of the identified indicators are relevant or new ones must be introduced. Similarly, more information on the supply chain might be desirable, particularly for transition strategies that go beyond climate risks and consider broader nature-related risks. While this is possible, and the tool can answer questions that are different from the pre-defined ones, the quality of answers needs to be more carefully checked as these new aspects would not have gone through our internal human evaluation process.

Methods

RAG Setup

To implement the RAG system, we use the LLama-Index software package. For the pipeline, we had to set a variety of parameters. Table 3 gives an overview of the RAG parameters.

For prompting the LLM, we rely on prior literature at the intersection of NLP and climate change[29]. The final RAG prompt requires basic information about the underlying company, which is directly retrieved with the prompt template shown in Figure 6. After retrieving the relevant information from the underlying report, the final prompt provided to the model is displayed in Figure 7 for a general question.

Parameter	Value
Chunk Size	350
Chunk Overlap	50
Top K Retrieval	8
Embedding	text-embedding-ada-002
Model	gpt-4-1106-preview
Answer Length	200

Table 3: RAG Parameters

As shown in Table 3, we prompt GPT-4 to obtain an answer as this model is the current state-of-the-art in evidence-based question-answering[28].

```
You are tasked with the role of a climate
    scientist and assigned to analyze a company's
     sustainability report. Based on the
    following extracted parts from the
    sustainability report, answer the given
    QUESTIONS.
If you don't know the answer, just say that you
    don't know by answering "NA". Don't try to
    make up an answer.

Given are the following sources:
---------------------- [BEGIN OF SOURCES]\n
{sources}\n
---------------------- [END OF SOURCES]\n

QUESTIONS:
1. What is the company of the report?
2. What sector does the company belong to?
3. Where is the company located?

Format your answers in JSON format with the
    following keys: COMPANY_NAME and
    COMPANY_SECTOR COMPANY_LOCATION.
Your FINAL_ANSWER in JSON (ensure there's no
    format error):
```

Figure 6: Prompt template to obtain basic information about the company.

```
You are a senior sustainability analyst with
    expertise in climate science evaluating a
    company's climate-related transition plan and
     strategy.

This is basic information to the company:
{basic_info}

You are presented with the following sources from
    the company's report:
---------------------- [BEGIN OF SOURCES]\n
{sources}\n
---------------------- [END OF SOURCES]\n

Given the source information and no prior
    knowledge, your main task is to respond to
    the posed question encapsulated in "||".
Question: ||{question}||

Please consider the following additional
    explanation to the question encapsulated in "
    +++++" as crucial for answering the question:
+++++ [BEGIN OF EXPLANATION]
{explanation}
+++++ [END OF EXPLANATION]

Please enforce the following guidelines in your
    answer:
1. Your response must be precise, thorough, and
    grounded on specific extracts from the report
     to verify its authenticity.
2. If you are unsure, simply acknowledge the lack
    of knowledge, rather than fabricating an
    answer.
3. Keep your ANSWER within {answer_length} words.
4. Be skeptical to the information disclosed in
    the report as there might be greenwashing (
    exaggerating the firm's environmental
    responsibility). Always answer in a critical
    tone.
5. Cheap talks are statements that are costless to
     make and may not necessarily reflect the
    true intentions or future actions of the
    company. Be critical for all cheap talks you
    discovered in the report.
6. Always acknowledge that the information
    provided is representing the company's view
    based on its report.
7. Scrutinize whether the report is grounded in
    quantifiable, concrete data or vague,
    unverifiable statements, and communicate your
     findings.
8. Start your answer with a "[[YES]]"" or ""[[NO]]
    "" depending on whether you would answer the
    question with a yes or no. Always compliment
    your judgment on yes or no with a short
    explanation that summarizes the sources in an
     informative way, i.e. provide details.

Format your answer in JSON format with the two
    keys: ANSWER (this should contain your answer
     string without sources), and SOURCES (this
    should be a list of the SOURCE numbers that
    were referenced in your answer).
Your FINAL_ANSWER in JSON (ensure there's no
    format error):
```

Figure 7: RAG prompt template enforcing structured output.

References

[1] Simon Dikau, Nick Robins, Agnieszka Smoleńska, Jens van't Klooster, and Ulrich Volz. Net zero transition plans: a supervisory playbook for prudential authorities. 2022.

[2] Ludovic Rappe, Noha Haddad, and Hervé Jeanmart. Credibility analysis of the company's transition plan. 2023.

[3] José Luis Reséndiz and Gireesh Shrimali. Assessing the credibility of climate transition plans in the aviation sector. Technical report, 2023.

[4] Xiaoyan Zhou and Gireesh Shrimali. Assessing the credibility of climate transition plans in the power sector. Technical report, 2023.

[5] Keith Jin Deng Chan, Bon Cheung, and Louis Yue Shen. How to assess the credibility of corporate transition pathways? application of the tcfd framework based on economic theory. *Application of the TCFD Framework based on Economic Theory (July 8, 2023)*, 2023.

[6] Nepomuk Dunz, Asjad Naqvi, and Irene Monasterolo. Climate transition risk, climate sentiments, and financial stability in a stock-flow consistent approach. *Climate Sentiments, and Financial Stability in a Stock-Flow Consistent Approach (April 1, 2019)*, 2019.

[7] Stefano Battiston, Yannis Dafermos, and Irene Monasterolo. Climate risks and financial stability, 2021.

[8] Francesca Diluiso, Barbara Annicchiarico, Matthias Kalkuhl, and Jan C Minx. Climate actions and macro-financial stability: The role of central banks. *Journal of Environmental Economics and Management*, 110:102548, 2021.

[9] Ran Duchin, Janet Gao, and Qiping Xu. Sustainability or greenwashing: Evidence from the asset market for industrial pollution. *Available at SSRN 4095885*, 2022.

[10] Louis Daumas. Financial stability, stranded assets and the low-carbon transition–a critical review of the theoretical and applied literatures. *Journal of Economic Surveys*, 2023.

[11] Elise Gourier and Hélène Mathurin. A greenwashing index. *Available at SSRN*, 2024.

[12] Niklas Höhne, Matthew Gidden, Michel den Elzen, Frederic Hans, Claire Fyson, Andreas Geiges, Louise Jeffery, Sofia Gonzales-Zuñiga, Silke Mooldijk, William Hare, and Joeri Rogelj. Wave of net zero emission targets opens window to meeting the paris agreement. *Nature Climate Change*, 11:820–822, 2021.

[13] Alexandra Luccioni, Emily Baylor, and Nicolas Duchene. Analyzing sustainability reports using natural language processing, 2020.

[14] Julia Anna Bingler, Mathias Kraus, Markus Leippold, and Nicolas Webersinke. How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk. *Swiss Finance Institute Research Paper,(22-01)*, 2023.

[15] Jingwei Ni, Julia Bingler, Chiara Colesanti-Senni, Mathias Kraus, Glen Gostlow, Tobias Schimanski, Dominik Stammbach, Saeid Ashraf Vaghefi, Qian Wang, Nicolas Webersinke, et al. Chatreport: Democratizing sustainability disclosure analysis through llm-based tools. *arXiv preprint arXiv:2307.15770*, 2023.

[16] Tobias Schimanski, Julia Bingler, Camilla Hyslop, Mathias Kraus, and Markus Leippold. Climatebert-netzero: Detecting and assessing net zero and reduction targets. *arXiv preprint arXiv:2310.08096*, 2023.

[17] Malte Toetzke, Anna Stünzi, and Florian Egli. Consistent and replicable estimation of bilateral climate finance. *Nature Climate Change*, 12:897 – 900, 2022.

[18] Anne J. Sietsma, James D. Ford, and Jan C. Minx. The next generation of machine learning for tracking adaptation texts. *Nature Climate Change*, 2023.

[19] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[20] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.

[21] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 2023.

[22] OpenAI. Gpt-4 technical report, 2023. URL https://cdn.openai.com/papers/gpt-4.pdf.

[23] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. *ACMComputing Surveys*, 55(12): 1–38, mar 2023. doi: 10.1145/3571730.

[24] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 9459–9474. Curran Associates, Inc., 2020.

[25] Tianjun Zhang, Shishir G. Patil, Naman Jain, Sheng Shen, Matei Zaharia, Ion Stoica, and Joseph E. Gonzalez. Raft: Adapting language model to domain specific rag, 2024.

[26] Jon Saad-Falcon, Omar Khattab, Christopher Potts, and Matei Zaharia. Ares: An automated evaluation framework for retrieval-augmented generation systems, 2024.

[27] Akari Asai, Zeqiu Wu, Yizhong Wang, Avirup Sil, and Hannaneh Hajishirzi. Self-rag: Learning to retrieve, generate, and critique through self-reflection, 2023.

[28] Tobias Schimanski, Jingwei Ni, Mathias Kraus, Elliott Ash, and Markus Leippold. Towards faithful and robust llm specialists for evidence-based question-answering, 2024.

[29] Jingwei Ni, Julia Bingler, Chiara Colesanti-Senni, Mathias Kraus, Glen Gostlow, Tobias Schimanski, Dominik Stammbach, Saeid Ashraf Vaghefi, Qian Wang, Nicolas Webersinke, Tobias Wekhof, Tingyu Yu, and Markus Leippold. CHATREPORT: Democratizing Sustainability Disclosure Analysis through LLM-based Tools. In Yansong Feng and Els Lefever, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 21–51, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-demo.3.

[30] Markus Leippold, Saeid Ashraf Vaghefi, Dominik Stammbach, Veruska Muccione, Julia Bingler, Jingwei Ni, Chiara Colesanti-Senni, Tobias Wekhof, Tobias Schimanski, Glen Gostlow, Tingyu Yu, Juerg Luterbacher, and Christian Huggel. Automated fact-checking of climate change claims with large language models, 2024.

[31] Europäische Zentralbank. Walking the talk: Banks gearing up to manage risks from climate change and environmental degradation results of the 2022 thematic review on climate-related and environmental risks publications office, 2022.

[32] Shuang Chen. Green investors and green transition efforts: Talk the talk or walk the walk? *Available at SSRN 4254894*, 2022.

[33] Julia Bingler, Chiara Colesanti Senni, Tobias Schimanski, and Markus Leippold. Net zero transition plans: Red flag indicators to assess inconsistencies and greenwashing. 2023.

Acknowledgements

This paper has received funding from the Swiss National Science Foundation (SNSF) under the project 'How sustainable is sustainable finance? Impact evaluation and automated greenwashing detection' (Grant Agreement No. 100018_207800), as well as partial funding from WWF Switzerland.

Supplementary Information

S.1 Advisory board

For the development of the conceptual framework, the selection of the indicators and for pilot testing the tool, we rely on the knowledge of more than 50 external experts that were part of our advisory board and pilot study group. Although we cannot disclose the names of the members, the advisory board includes representatives from financial supervision, central banks, governmental organizations, NGOs, and industry practitioners.

S.2 Reviewed frameworks

The frameworks reviewed to identify the final list of indicators are displayed in Table S.1[33]. For each framework, the criteria used to assess transition disclosures were extracted and combined to identify a common ground.

Initiative	Year	Preparer	Focus	Assessment
ACT	2021	Corporates	Strategy	Ambition, credibility, feasibility
CSLN	2021	Financial institutions	Strategy	Disclosure, ambition, credibility
TCFD	2021	Corporates	Risk	Disclosure
UNEP-FI	2021	Financial institutions	Strategy	Ambition, credibility
WBA	2021	Corporates	Strategy	Feasibility
CPI	2022	Corporates	Strategy	Credibility
ESRS	2022	Corporates	Risk	Disclosure
GFANZ NZTP	2022	Financial institutions	Strategy	Disclosure, ambition, feasibility
GFANZ RETP	2022	Real economy corporates	Strategy	Disclosure
IFRS ISSB	2022	Corporates	Risk	Disclosure
NewClimate et al	2022	Corporates	Strategy	Disclosure, credibility
R2Z	2022	Corporates	Strategy	Ambition
SBTi FINZ	2022	Financial institutions	Strategy	Feasibility
TPI	2022	Corporates	Strategy	Ambition
TPT	2022	Corporates	Strategy	Disclosure
UN HLEG	2022	Corporates	Strategy	Ambition
WWF	2022	Corporates	Strategy	Credibility
CA100+	2023	High emitting corporates	Strategy	Disclosure, ambition
CBI CBS4	2023	Real economy corporates	Strategy	Disclosure, ambition, credibility
CDP	2023	Corporates	Strategy	Disclosure, ambition, credibility
IIGCC	2023	Corporates	Strategy	Credibility
NGFS	2023	Corporates, fin. institutions	Strategy, risk	Credibility
NZAOA	2023	Financial institutions	Strategy	Feasibility
OxSFG	2023	Real economy corporates	Strategy	Credibility
PwC et al	2023	Corporates	Strategy	Feasibility
RI	2023	Financial institutions	Strategy	Feasibility
SBTi Net Zero	2023	Corporates	Strategy	Ambition
WWF PtP	2023	Real economy corporates	Strategy	Ambition, feasibility

Table S.1: Transition plan frameworks assessed for the proposed credibility, ambition, and feasibility assessment framework.

S.3 Assessment indicators

Tables S.2-S.5 display the detailed indicators and their classification along two dimensions. First, we build on the widely known classification of companies' activities and merge them into Target, Governance, Strategy, and Tracking categories (building, e.g., on the initial classification by the TCFD, which has further evolved in the context of transition plan assessments). While this is of secondary importance for the underlying paper, it can help end users in their application cases. Second, we classify the indicators according to "Walk" and "Talk". Overall, we classify 34 of the 64 indicators as "Walk" and 30 as "Talk".

Identifier	Question	Walk or Talk
Target		
1	Does the company report an absolute GHG emission reduction target for the company?	T
2	If the company communicates GHG emission intensity targets, does the company show that the company's intensity targets are in line with its absolute emission targets?	T
3	Does the company report a company-wide net zero GHG emissions target?	T
4	Does the company state explicitly that it plans to achieve its net zero target until 2040 or 2050 at the latest?	T
5	Does the company state explicitly that it plans to cut its absolute GHG emissions by 50% (by half) until 2030 at the latest?	T
6	Does the company report its absolute emission target for the company's scope 1 GHG emissions?	T
7	Does the company report its absolute emission target for the company's scope 2 GHG emissions?	T
8	Does the company report its absolute emission target for the company's scope 3 GHG emissions?	T
9	Does the company report its GHG emission reduction interim targets for achieving the overall goal?	T
10	Does the company state explicitly that the interim targets are in line with specific 1.5-degree orderly sector transition pathways, which are based on front-loaded activities and no or limited emission overshoot?	T
11	If carbon credits and offsets are reported to be used by the company, does the company explicitly state that it will use them exclusively for residual unabatable emissions or beyond value chain mitigation support?	T
12	If carbon credits and offsets are reported to be used by the company, does the company state explicitly that carbon credits and offsets will be only used when the company can ensure that the emission reduction or emission avoidance is sustained permanently?	T

Table S.2: Final list of Target indicators selected and their classification.

Identifier	Question	Walk or Talk
Governance		
13	Does the company explain its governance structure for managing the climate transition?	T
14	Does the company explain how it ensures that the board members have the required skills to sign off and oversee the climate transition plan implementation?	T
15	Does the company report its available in-house skills and additional capacity needed to implement the climate transition plan?	T
16	Does the company report a strategy on how it aims to fill the additional skill and capacity needs required to implement its climate transition plan?	T
17	Does the company report how its board oversees the climate transition plan implementation?	T
18	Does the company report that it ensures that the company's board is informed at least quarterly about the progress against achieving the climate transition plan targets?	W
19	Does the company provide a higher share of remuneration and bonuses that are linked to the successful implementation of the climate transition plan interim targets compared to the general part of variable compensation for executives and managers?	W
20	Does the company report that the climate transition plan targets and information contained in the report have been subject to external assurance and validation?	T
21	Does the company state explicitly that it uses the same organizational boundaries for setting and achieving its climate targets as it does for financial accounting?	T

Table S.3: Final list of Governance indicators selected and their classification.

Identifier	Question	Walk or Talk
Strategy		
22	Does the company provide comprehensive evidence that it fully and completely integrates its climate strategy into its business strategy, product development, operations, financial and human resources, asset management, and asset decommissioning?	W
23	Does the company report quantitative or quantifiable sub-targets in line with their climate targets and their climate key performance indicators?	T
24	Does the company report the use of scenario envelopes to set targets and perform sensitivity analysis?	T
25	Has the company reported the key assumptions that form the basis of its transition plan?	T
26	Does the company report a renewable energy strategy and activities covering renewable energy build-out, procurement, and consumption?	W
27	Does the company report a strategy and activities for the expansion of and investments in climate solutions and climate solution technologies?	W
28	Does the company report its opex planning to ensure it meets its climate interim targets?	W
29	Does the company report its capex planning to ensure it meets its climate interim targets?	W
30	Does the company report its strategy and activities towards net zero aligned (or green) revenues?	W
31	Does the company report its strategy and activities to align all its Research and Development (R&D) activities with net zero targets?	W
32	Does the company report a strategy with specific activities and metrics of success for net zero engagement with its upstream value chain, i.e., with its suppliers?	T
33	Does the company report a strategy with specific activities and metrics of success for net zero engagement with its downstream value chain, i.e., with its customers or investees?	T
34	Does the company report an engagement strategy and activities with specific climate policies that it directly advocates for with policymakers?	W
35	Does the company report a strategy and activities with specific activities and engagement goals for net zero policies with industry associations?	W
36	Does the company report serious consequences and escalation strategies if net zero engagement is ineffective at upstream, downstream, policymaker, and industry association levels?	W
37	Does the company state explicitly that it stopped or will immediately stop any support or activities in new additional fossil fuel exploration and extension of fossil fuel supply?	W
38	Does the company report a strategy and activities for the decommissioning and canceling of planned or existing fossil fuel exploration and supply infrastructure?	W
39	Does the company report a strategy and activities to phase out its use and support of fossil fuel-consuming products and technologies?	W
40	Does the company report a strategy for a just transition, including monitoring and activities to mitigate adverse impacts of the net zero transition on its own workforce and indirectly affected workers and local communities?	T
41	Does the company report that it develops specific just transition plans with its own climate transition-affected workforce, local communities, and relevant stakeholders?	T
42	Does the company report a strategy and activities to mitigate adverse impacts on the natural environment and the provision of ecosystem services?	W
43	Does the company report a strategy and activities to halt and reverse deforestation by 2025?	W
44	Does the company report a strategy and activities to halt and reverse biodiversity loss by 2030?	W
45	Does the company report a strategy and activities to significantly reduce water consumption and pollution?	W

Table S.4: Final list of Strategy indicators selected and their classification.

Identifier	Question	Walk or Talk
Tracking		
46	Does the company report its scope 1 GHG emissions for the past year?	W
47	Does the company report its scope 2 GHG emissions for the past year?	W
48	Does the company report its scope 3 GHG emissions for the past year?	W
49	Does the company report the coverage of scope 3 categories included and the reasons for the exclusion of categories?	T
50	Does the company report its annual progress of reducing GHG emissions to achieve its emission reduction or net zero targets?	W
51	Does the company report its absolute scope 1 GHG emissions for the past 5 years?	W
52	Does the company report its absolute scope 2 GHG emissions for the past 5 years?	W
53	Does the company report its absolute scope 3 GHG emissions for the past 5 years?	W
54	Does the company report a decline in its scope 1 GHG emission intensity for the past 5 years?	W
55	Does the company report a decline in its scope 2 GHG intensity for the past 5 years?	W
56	Does the company report a decline in its scope 3 GHG intensity for the past 5 years?	W
57	Does the company report the specific drivers and reasons for the company's observed actual GHG emission changes?	T
58	Does the company report annual progress against its deforestation targets?	W
59	Does the company report the amount of climate-aligned capex that supports its net zero transition?	W
60	Does the company report the amount of climate transition misaligned capex?	W
61	Does the company report the amount of climate transition-aligned revenues that support the global net zero transition?	W
62	Does the company report the amount of climate transition misaligned revenues?	W
63	Does the company assess and report the alignment of its transition plan with its policy positions and its trade association's policy positions and lobbying?	T
64	Does the company report its engagement activities with the companies it invests in its own financial portfolio (including voting and proxy voting) undertaken in the relevant reporting period?	T

Table S.5: Final list of Tracking indicators selected and their classification.

S.4 Disclosed indicators

In Figures S.1-S.4, we depict the number of disclosed indicators in the four different categories (Target, Governance, Strategy, and Tracking) for the CA100+ companies. We see that indicators are disclosed across all categories and that there is a wide divergence in the coverage of indicators within each category. As highlighted in the main text, the main difference in the disclosures relates to walk and talk indicators. The most and least disclosed indicators are also described and discussed in the main text.

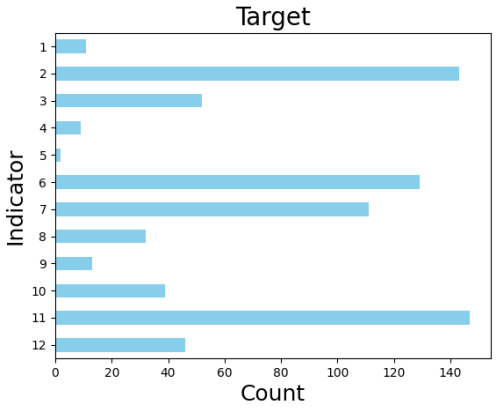


Figure S.1: Disclosed indicators in the Target group.

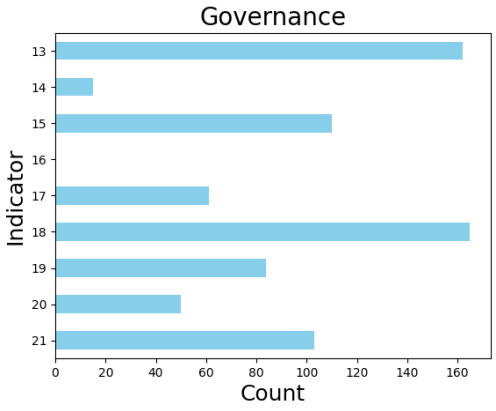


Figure S.2: Disclosed indicators in the Governance group.

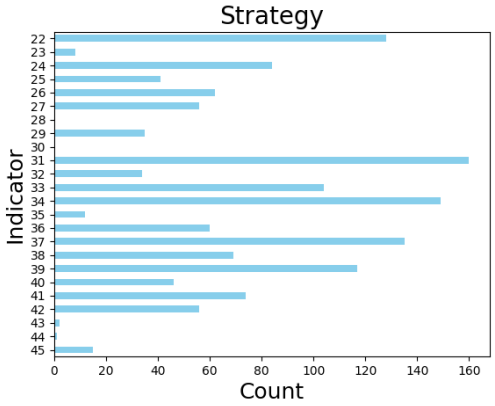


Figure S.3: Disclosed indicators in the Strategy group.

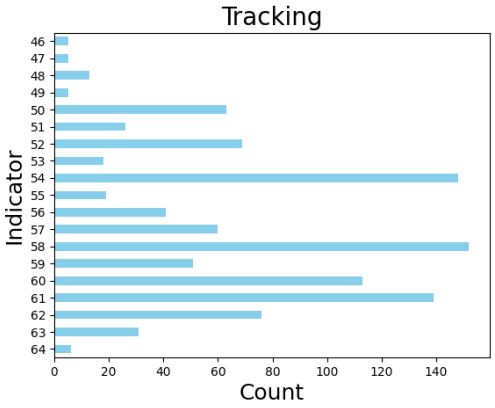


Figure S.4: Disclosed indicators in the Tracking group.

S.5 Expert-Centric Question Extension

The analysis of the answers initially provided by the tool revealed one particular disadvantage of generic RAG systems. Individual questions might be too vague, and the model's interpretation of the question might lead to an unsatisfactory outcome. Hence, we extended the prompt provided to the LLM with expert-centric knowledge and details. In essence, these extensions provide guidelines on how the model should answer the questions and the information that it should take into account when assessing the available information. They can contain restrictions, focus areas of the question, or extended explanations of some concepts. They are the result of an iterative process analyzing several responses of the model for each individual indicator. In this process, we build on the distinctive advantage of LLMs to process expert feedback and incorporate it into their answers.

Figure S.5 offers an example of such guidelines. It shows that we ask the model to follow specific restrictions. In most cases, the guidelines require the model to be more conservative or focus on specific aspects.

```
QUESTION: Does the company state explicitly that
    it plans to achieve its net zero target by
    2040 or 2050 at the latest?

QUESTION EXTENSION: Answer with "YES" if this is
    the case and "NO" otherwise. Summarize the
    timeframe for which the net zero target is to
     be achieved if this information is stated.
    Otherwise, if the company, in general, does
    not have a net zero target, state that the
    company has no net zero targets, and
    therefore, it also has no goal to achieve net
     zero by 2040 or 2050 at the latest. If the
    company has a net zero target but no
    timeframe for the achievement is reported,
    report the target and state that there is no
    year reported until the company aims to
    achieve its net zero targets. Focus on
    summarizing the relevant information, and do
    not judge the ambition, feasibility, or
    credibility of the target. Do not reference
    the sources and the chunks in the summary
    answer.
```

Figure S.5: Example of a question guideline.

The complete list of questions and extensions can be found in Tables S.6-S.9.

S.6 Answer Explanation and Sources Reported

While the dominant focus of this work lies on exploring the (non-)disclosure of the assessment indicators, the NLP tool also produces two important pieces of information: an answer explanation and the sources of the information used in the answers. The answer explanation can provide the user with a deeper understanding of the verdict. Together with the sources, it can help the user validate or investigate the underlying report further. This considerably increases the transparency of the approach. Users can check whether the NLP tool truly bases its answers on actual information available in the disclosure and accurately represents information or can use the tool to ease their search process for context and further information within the document itself.

Figure S.6 shows a typical answer. The first piece of information is the final verdict. Although the company does not disclose a target for reducing scope 3 GHG emissions, it discloses information about recycling materials or reducing their battery carbon footprint. This level of detail, combined with the sources, returns a more holistic assessment of the indicators.

```
QUESTION: Does the company report its absolute
    emission target for the company's scope 3 GHG
     emissions?

ANSWER: [[NO]] The company does not explicitly
    report its absolute emission reduction
    targets for scope 3 GHG emissions. The
    sources provided discuss various strategies
    and targets related to reducing the company's
     carbon footprint, such as reducing battery
    carbon footprint, increasing the use of
    recycled materials, and reducing emissions in
     their parts and materials supply chain.
    However, these targets are not explicitly
    categorized under Scope 3 emissions, nor is
    there a clear absolute emission reduction
    target for Scope 3 emissions. The targets
    mentioned are either related to scope 1 and 2
     emissions, such as emissions at their sites,
     or are not clearly defined in terms of scope
     categorization. Therefore, no explicit Scope
     3 absolute emission reduction target could
    be identified.

SOURCE PAGES (ordered by importance): 21, 20, 19
```

Figure S.6: Example of an answer explanation and sources reported.

Question	Expert-Centric Question Extension
Target	
Does the company report an absolute GHG emission reduction target for the company?	Focus only on absolute emission reduction targets. Be aware that a net zero target or a carbon neutrality target can be counted as an absolute emission reduction target. If only a net zero target or a carbon neutrality target is reported, state that it is counted as an absolute emission reduction target in the given analysis but that the effective absolute reduction in emissions needs to be carefully assessed. Also, be aware that statements "reducing emissions to return to 2020 levels" are also absolute emission reduction targets, although they are indirectly formulated. Answer with "YES" if the company reports an absolute emission target and with "NO" otherwise. Be aware that intensity targets are not absolute emission reduction targets. Intensity targets are, for example, efficiency targets and targets that aim to reduce the amount of emissions relative to another economic or physical metric (for example dollars, or kilometers driven). If only intensity targets are mentioned in the report, add them to your answer, but state that these are not absolute emission reduction targets. Do not include past targets that have been achieved in your response. Do not provide a summarizing sentence at the end of your response, and do not speculate whether a certain target is an indicator for climate action or supportive of any other climate goals.
If the company communicates GHG emission intensity targets, does the company show that the company's intensity targets are in line with its absolute emission targets?	Focus on GHG emission intensity targets. If no intensity targets are communicated, state that you did not find GHG emission intensity targets in the document. If no intensity targets are communicated, state "N/A". If intensity targets are communicated, state "YES" if they are shown to be in line with the absolute emission intensity targets. If no absolute emission targets are reported, or if the intensity targets are not shown to be in line with the absolute emission targets, state "NO". If there are other intensity targets reported, for example, energy intensity targets or sectoral materials-related intensity targets, summarize them in your response, but do not consider them for the overall Yes/No assessment.
Does the company report a company-wide net zero GHG emissions target?	Focus on the simple presence of a net zero target for the reporting company. Do not include other targets, strategies, or activities that the company undertakes. If the company is part of a net zero alliance, you mention it but clearly state that you did not identify an own net zero target for the company and that a clear company-wide commitment to net zero is missing in the report. Do not judge the activities or the target in terms of greenwashing, climate goal alignment, or seriousness. Just summarize the information you find on net zero targets and net zero commitments. Be aware that carbon neutrality goals are also net-zero targets. If the company has net zero targets for parts of its operations or supports net zero targets in its value chain but does not have its own net zero target, summarize the targets but state that it is not a target from the company. State "YES" if a company-wide net zero or carbon neutrality target is reported. State "NO" if no net zero or carbon neutrality target is reported.
Does the company state explicitly that it plans to achieve its net zero target by 2040 or 2050 at the latest?	Answer with "YES" if this is the case, and "NO" otherwise. Summarize until which timeframe the net zero target is to be achieved if this information is stated. Otherwise, if the company in general does not have a net zero target, state that the company has no net zero target, and therefore, it also has no goal to achieve net zero by 2040 or 2050 at the latest. If the company has a net zero target, but no timeframe for the achievement is reported, report the target and state that there is no year reported until when the company aims to achieve its net zero targets. Focus on summarizing the relevant information, and do not judge the ambition, feasibility, or credibility of the target. Do not reference the sources and the chunks in the summary answer.

Question	Expert-Centric Question Extension
Target	
Does the company state explicitly that it plans to cut its absolute GHG emissions by 50% (by half) until 2030 at the latest?	No additional guidelines
Does the company report its absolute emission target for the company's scope 1 GHG emissions?	Bear in mind that scope 1 emissions are any direct GHG emissions from a company, for example emitted in its production processes, or via direct heat and energy production on site by the company. State "YES" if there are absolute emission reduction targets reported for scope 2 emissions of the company. State "NO" otherwise. If there is just a combined scope 1 and scope 2 target reported, state this in your summary, but clarify that it is a combined target, and therefore, no explicit scope 1 absolute emission reduction target could be identified.
Does the company report its absolute emission target for the company's scope 2 GHG emissions?	Bear in mind that scope 2 emissions are any indirect GHG emissions from a company's purchased energy for electricity, heating, or cooling. These emissions are usually operation-related emissions. State "YES" if there are absolute emission reduction targets reported for scope 2 emissions of the company. State "NO" otherwise. If there is just a combined scope 1 and scope 2 target reported, state this in your summary, but clarify that it is a combined target, and therefore, no explicit scope 2 absolute emission reduction target could be identified.
Does the company report its absolute emission target for the company's scope 3 GHG emissions?	Bear in mind that scope 3 emissions are any GHG emissions in a company's value chain, beyond its direct operational control. This includes upstream and downstream activities and processes. It mainly includes the use of company's products, the emissions of financed projects and businesses beyond its own operations, the emissions embodied in its purchased goods and services, the emissions from business travels, and others. State "YES" if there are absolute emission reduction targets reported for scope 3 emissions of the company. State "NO" otherwise. If there is just a combined scope 1, scope 2 and scope 3 target reported, state this in your summary, but clarify that it is a combined target, and therefore, no explicit scope 3 absolute emission reduction target could be identified. Do not judge the presence or absence of the scope 3 emissions target. Do not provide information on the sources and chunks in the written summary.
Does the company report its GHG emission reduction interim targets for achieving the overall goal?	Bear in mind that GHG emission reduction interim targets are targets that the company aims to achieve to ensure it achieves its overall emission target. Interim targets provide a way to measure and evaluate the effectiveness of mitigation efforts and can enhance the credibility of long-term goals. If interim goals are reported, state "YES". State "NO" otherwise. Do not judge the ambition or credibility of the interim goals.
Does the company state explicitly that the interim targets are in line with specific 1.5 degrees orderly sector transition pathways, which are based on frontloaded activities and no or limited emission overshoot?	No additional guidelines

Question	Expert-Centric Question Extension
Target	
If carbon credits and offsets are reported to be used by the company, does the company explicitly state that it will use them exclusively for residual unabatable emissions or beyond value chain mitigation support?	Answer with "YES" if this is the case. Answer with "NO" if the company report mentions the use of carbon offsets or carbon credits, but does not include an explicit statement that the company will use them exclusively for residual unabatable emissions or beyond value chain mitigation support. Answer with "N/A" if the company does not plan to use carbon offsets, or if the information about the use of carbon credits and carbon offsets is not available. Justify in the summary why you came to the conclusion that YES, NO or N/A are the correct answers to the question.
If carbon credits and offsets are reported to be used by the company, does the company state explicitly that carbon credits and offsets will be only used when the company can ensure that the emission reduction or emission avoidance is sustained permanently?	Answer with "YES" if this is the case. Answer with "NO" if the company report mentions the use of carbon offsets or carbon credits, but does not include an explicit statement that the company will use them only when the company can ensure that the emission reduction or emission avoidance is sustained permanently. Answer with "N/A" if the company does not plan to use carbon offsets, or if the information about the use of carbon credits and carbon offsets is not available. Justify in the summary why you came to the conclusion that YES, NO or N/A are the correct answers to the question.

Table S.6: Expert-centric question extensions for Target questions.

Question	Expert-Centric Question Extension
Governance	
Does the company explain its governance structure for managing the climate transition?	Answer with "YES" if information is available on for example board-level committees with climate responsibilities, a climate representative at/or reporting to the executive/board level, a clear team responsible for climate projects, reporting and disclosures. Answer with "NO" if there is no information about the company's governance structure for the climate transition.
Does the company explain how it ensures that the board members have the required skills to sign off and oversee the climate transition plan implementation?	No additional guidelines
Does the company report its available inhouse skills and additional capacity needs to implement the climate transition plan?	No additional guidelines
Does the company report a strategy on how it aims to fill the additional skill and capacity needs required to implement its climate transition plan?	No additional guidelines
Does the company report how its board oversees the climate transition plan implementation?	No additional guidelines
Does the company report that it ensures that the company's board is informed at least quaterly about the progress against achieving the climate transition plan targets?	No additional guidelines
Does the company provide a higher share of remuneration and bonuses that are linked to the successful implementation of the climate transition plan interim targets compared to the general part of variable compensation for executives and managers?	Provide specific examples of executive management remuneration linked to progress towards achievement of transition plan interim targets. Specify how the percentage linked to the progress towards and achievement of transition plan interim targets compares to the variable compensation in general. Answer with "YES" if the share of climate-related variable compensation compared to non-climate-related variable compensation is higher. If there is only information available about variable remuneration for climate targets in general, or if the share of climate-related variable compensation is not higher than the overall variable compensation, state this information and answer with "NO."
Does the company report that the climate transition plan targets and information contained in the report have been subject to external assurance and validation?	Provide specific examples of defined assurance and verification levels of the transition plan and statements by third parties. Focus on climate assurance and verification only. Consider Second Party Opinions (SPOs) as verification by a third party. Target validation by the Science-based targets initiative should be counted as an external validation of the transition plan targets. If a third party validation is reported, or the organisation that provides the third party valiation is stated, answer with "YES". If the report only states information about reporting initiatives like CDP, TCFD and GRI, state "NO". Do not include reporting and disclosure frameworks and voluntary initiatives in the answer.
Does the company state explicitly that it uses the same organizational boundaries for setting and achieving its climate targets as it does for financial accounting?	No additional guidelines

Table S.7: Expert-centric question extensions for Governance questions.

Question	Expert-Centric Question Extension
Strategy	
Does the company provide comprehensive evidence that it fully and completely integrates its climate strategy into its business strategy, product development, operations, financial and human resources, asset management, and asset decommissioning?	Provide specific examples of how the company's climate strategy is integrated into its business strategy, product development, operations, financial and human resources, asset management, and asset decommissioning. Focus on aspects related to the climate transition, and not on other sustainability or nature-related topics. Do not rephrase the company's climate or sustainability targets. Instead, look for information that shows how the company implements these targets into the core of the company's strategy, activities and management. If you find this information, summarize it and answer with "YES". Do not judge whether the activities are sufficient. If you find only partial information, summarize the information, answer with "NO" and highlight that additional information would be required to assess whether the company fully integrates its climate strategy across its activities.
Does the company report quantitative or quantifiable subtargets in line with their climate targets and their climate key performance indicators?	Provide information about the specific quantifiable subtargets that the company has identified to achieve their climate targets and climate key performance indicators, the time scale (if provided), and scope. State explicitly if the time scale or the application scope of the subtarget is missing. Answer "YES" if you find quantitative or quantifyable subtargets, and state the targets explicitly. Answer "NO" otherwise.
Does the company report the use of scenario envelopes to set targets and perform sensitivity analysis?	Provide specific examples of the scenarios, model ensembles and scenario envelopes used by the company to set targets and perform targets and pathways sensitivity analysis. Also include the time scale and scope of the scenario analyses. State whether the company makes reference to model constraints and whether it is aware of the limitations of modelling, and the need for resilience planning. If you find information about the use of multiple scenarios, model families or scenario envelopes directly linked to the targets and the plans to implement the targets, answer "YES". If you only find such information in terms of how the company assesses its climate risk exposure, answer "NO" and state that the company uses multiple scenarios, but it does not seem to use scenario enveloped to set its climate targets and identify a resilient strategy to achieve the targets.
Has the company reported its key assumptions that form the basis of its transition plan?	Provide specific examples of the strategic assumptions that the company reports as basis of its transition plans. These could include for instance assumptions about the development of consumer preferences, input prices, sector policies, economic development, and others. Answer "YES" if you find information about the assumptions underlying the transition plan. Answer "NO" if you do not find this information.
Does the company report a renewable energy strategy and activities, covering renewable energy build out, procurement and consumption?	Provide specific examples of how the company plans to increase renewable energy build out, procurement and consumption. If you find such examples and ideally a strategy, answer "YES". If the company does not report activities to expand renewable energies, answer "NO". Focus only on specific renewable energy activities, and do not include general net zero emission targets or further activities of the company for climate action in this answer.
Does the company report a strategy and activities for the expansion of and investments in climate solutions and climate solution technologies?	Focus on strategies that focus on forward-looking and future investment in specific climate solutions. This includes, but is not limited to, research and development (R&D), investing in early stage climate solutions, or acquisitions and substantial capital provision to scale up the climate solutions branch of the company's business. Do not repeat what the company has been doing in the past. Do not include general statements about the climate targets of the company. If you find specific information about how the company supports the expansion of climate solution technologies, answer "YES" and summarize the specific activities it undertakes, the specific technologies it focuses on, the interim targets it identifies, and the associated timeframe. If you do not find information about the specific climate solutions support by the company, answer "NO" and explain that there is no information future investmens or R&D to support the build-out of climate solutions available.
Does the company report its opex planning to ensure it meets its climate interim targets?	Focus on quantified targets and achievements, including how the company plans to shift opex to be aligned with the climate targets, the climate strategy and the specific interim targets. If you find information about how the company aligns its opex with the climate targets, reply "YES". If available, provide the quantitative information about the amount and timeframe towards full alignment of opex with the climate targets. If you do not find opex information, or if opex information is available but it is not shown to be aligned with the strategies to achieve the climate targets, answer "NO".

Question	Expert-Centric Question Extension
Strategy	
Does the company report its capex planning to ensure it meets its climate interim targets?	Focus on quantified targets and achievements, including how the company plans to shift capex to be aligned with the climate targets, the climate strategy and the specific interim targets. If you find information about how the company aligns its capex with the climate targets, reply "YES". If available, provide the quantitative information about the amount and timeframe towards full alignment of capex with the climate targets. If you do not find capex information, or if capex information is available but it is not shown to be aligned with the strategies to achieve the climate targets, answer "NO".
Does the company report its strategy and activities towards net zero aligned (or green) revenues?	Focus on quantified targets and achievements for the company's own revenues, including its products and services, and how they plan to shift away from net zero emissions misaligned sources of revenues. If you only find a collection of business activities, but not a fully developed strategy on how to align all revenues of the company with the company's climate targets, answer "NO". If the company reports a plan to align all its revenues with its climate targets, state "YES". Do not include targets to align clients' revenues with the climate goals. Focus on the company's own revenue strategy. Do not include the cost-side of the profits, for example renewable energy purchases or energy efficiency. Focus on the climate alignment of the revenues that the company generates from its products and services.
Does the company report its strategy and activities to align all its Research and Development (R&D) activties with net zero targets?	Focus on strategies specific to Research and Development, do not summarize general net zero targets or general activties. If the company does not report on R&D expenditures, if it is not active in R&D, or if the R&D expenditures are not fully or only partially aligned with net zero targets, state "NO". If the company provides information how it aligns all its R&D activties with net zero targets, state "YES".
Does the company report a strategy with specific activities and metrics of success for net zero engagement with its upstream value chain, i.e. with its suppliers?	Provide specific information about the company's supplier engagement strategy, including actual engagement activities with its suppliers and metrics for success. Do not reference general net zero targets or general scope 1 scope 2 or scope 3 targets. Do not include customer engagements, focus on the company's suppliers. State "YES" if a specific strategy with specific activities and metrics for success is reported. State "NO" if the company does not report its upstream value chain strategy, or if the value chain engagement is not associated with specific activities.
Does the company report a strategy with specific activities and metrics of success for net zero engagement with its downstream value chain, i.e. with its customers or investees?	Provide specific information about the company's customer or investee engagement strategy, including actual engagement activities with its customers or investees and metrics for success. Do not reference general net zero targets or general scope 1 scope 2 or scope 3 targets. Do not include supplier engagements, focus on the company's customers or investees. State "YES" if a specific strategy with specific activities and metrics for success is reported. State "NO" if the company does not report its upstream value chain strategy, or if the value chain engagement is not associated with specific activities.
Does the company report an engagement strategy and activities with specific climate policies that it directly advocates for with policy makers?	Provide specific activities of policy maker engagement and the policies advocated for. Do not include general activties and engagements with clients or suppliers. Focus on active engagement with policy makers and political stakeholders. Include public speaking and public positioning for the climate policies needed. Do not include event attendance and indirect support via business organisations or business initiatives. If there is information about specific climate policies that the company advocates for with policy makers, state "YES". If there is no information about policy maker engagements, or if the information lacks information about specific climate policies, state "NO".
Does the company report a strategy and activities with specific activities and engagement goals for net zero policies with industry associations?	Identify specific activities with industry associations and industry groups where the company is a member, to align the industry associations' policy positions with the net zero transition. Provide specific information on where the company leads within an industry group, for example by chairing a committee or sitting on the steering committee. Provide information about the company's activities and specific policy goals that it advocates for. Do not include the company's own climate targets. Answer "YES" if a strategy with specific engagement goals and activities is defined. Answer "NO" if there is no information about engagements within the industry associations, or if the informatiom does not include specific activities and policy goals.

Question	Expert-Centric Question Extension
Strategy	
Does the company report serious consequences and escalation strategies if net zero engagement is ineffective at upstream, downstream, policy maker and industry association level?	Do not include general climate targets or the company's overall climate goals and activities. Do not reference the company's general climate engagement. Instead, focus on whether the company defines specific and serious escalation activities for the case of ineffective engagements. Provide specific examples for each level (upstream, downstream, policy makers and industry associations). Answer "YES" if you find specific information about consequences and escalation plans or activities of the company if its engagement with suppliers, customers, policy makers or industry associations is not successful. Include information about the measurement of unsuccessful, and the timeline for engagement until when the company uses its escalation activities. Examples of consequences and escalations for unsuccessful engagements are quitting contracts, no renewal of business relations, divesting, leaving an industry organisation, and others.
Does the company state explicitly that it stopped or will immediately stop any support or activities in new additional fossil fuel exploration and extention of fossil fuel supply?	Include and summarize statements for coal, oil and gas into the answer. If you only find a committment for coal, state this, but answer with "NO", and explain that it would need to include also a statement on oil and gas. If the company states explicitly that it already stopped or will stop supporting additional fossil fuel exploration or the extention of fossil fuel supply, answer "YES".
Does the company report a strategy and activities for the decommissioning and canceling of planned or existing fossil fuel exploration and supply infrastructure?	Provide specific information where available. Do not inclue landscape restoration activities for decommissioned sites, and the general decommission process. Reply "YES" if the company commits to decommission or cancel all existing or planned fossil fuel projects. Reply "NO" if the company only states that it has decommissioned some specific sites but there is no evidence of a structural committment that applies to all planned or existing fossil fuel infrastructure of the company.
Does the company report a strategy and activities to phase out its use and support of fossil fuel consuming products and technologies?	Specify whether the company has a defined strategy and time-bound goal for their product changes. Phase out means reducing these products to zero over time, highlight how the company sets interim targets towards this goal. Focus on end consumer products like combustion engine cars, and on business to business products like manufacturing equipment and technology. If you find comprehensive information about how the company aims to phase out its own use and the support of fossil fuel consuming products and technologies, ideally including a timeframe, state "YES". If you only find information about general policies of the company to reduce or partially halt the use and support of fossil fuel-consuming products and technologies, state "NO".
Does the company report a strategy for a just transition, including monitoring and activities to mitigate adverse impacts of the net zero transition on its own workforce and indirectly affected workers and local communities?	Do not include general climate targets of the company, and general activties. Focus on specific just transition strategies that the company defines, and how it monitors the effects of the transition on possibly affected own workers and indirectly affected workers and local communities. Reply "NO" if there is only general information about the importance of the transition for local communities available. Reply "YES" if the information is specific and showcases a well developed strategy to manage the transition in a fair and equitable manner.
Does the company report that it develops specific just transition plans with its own climate transition affected workforce, local communities and relevant stakeholders?	Do not include general climate targets of the company, and general activties. Focus on specific existing plans or plans under development that the company developed explicitly together with its affected own workers and indirectly affected workers and local communities. Reply "NO" if there is only general information about the importance of the transition for local communities available. Reply "YES" if the information is specific and showcases plans that have been developed with the affected parties to manage the transition in a fair and equitable manner.
Does the company report a strategy and activities to mitigate adverse impacts on the natural environment and the provision of ecosystem services?	Ensure that you only reply "YES" if there is a specific strategy with associated activities reported on how the company deals with adverse impacts on the natural environment and the provision of ecosystem services. If there is only information available about various nature initiatives or a loose collection of some project examples, summarize these activties, but reply "NO" and explain that a comprehensive strategy is not defined. Do not include general climate or ESG strategies and actvities of the company.
Does the company report a strategy and activities to halt and reverse deforestation by 2025?	Look for time-bound, specific goals, associated activities, and metrics to measure success. If there is a strategy available, reply "YES". If the company reports only some activities without a broader strategy that covers all its businesses, reply "NO".

Question	Expert-Centric Question Extension
Strategy	
Does the company report a strategy and activities to halt and reverse biodiversity loss by 2030?	Look for time-bound, specific goals, associated activities, and metrics to measure success. If there is a strategy available, reply "YES". If the company reports only some activities without a broader strategy that covers all its businesses, reply "NO".
Does the company report a strategy and activities to significantly reduce water consumption and pollution?	Look for time-bound, specific goals, associated activities, and metrics to measure success. If there is a strategy available, reply "YES". If the company reports only some activities without a broader strategy that covers all its businesses, reply "NO".

Table S.8: Expert-centric question extensions for Strategy questions

Question	Expert-Centric Question Extension
Tracking	
Does the company report its scope 1 GHG emissions for the past year?	Do not include general emission targets or climate goals of the company. Focus on the actual scope 1 GHG emissions that the company has emitted. Include the precise figures of the emitted scope 1 GHG emissions that the company reports. Reply "YES" if the precise quantitative information for the scope 1 GHG emissions is available. Reply "NO" otherwise. If available, include the information on the calculation approach or approaches that have been applied by the company.
Does the company report its scope 2 GHG emissions for the past year?	Do not include general emission targets or climate and renewable energy goals of the company. Focus on the actual scope 2 GHG emissions that the company has emitted. Include the precise figures of the emitted scope 2 GHG emissions that the company reports. Reply "YES" if the precise quantitative information for the scope 2 GHG emissions is available. Reply "NO" otherwise. If available, include the information on the calculation approach or approaches that have been applied by the company.
Does the company report its scope 3 GHG emissions for the past year?	Do not include general emission targets or general climate goals of the company. Focus on the actual scope 3 GHG emissions that the company has emitted. Include the precise figures of the emitted scope 3 GHG emissions that the company reports. Reply "YES" if the precise quantitative information for the scope 3 GHG emissions is available. Reply "NO" otherwise. If available, include the information on the calculation approach or approaches that have been applied by the company.
Does the company report the coverage of scope 3 categories included, and reasons for the exclusion of categories?	Focus on the coverage of the scope 3 emission categories as outlined by the GHG protocol, namely purchased goods and services, capital goods, fuel- and energy-related activities, transportation and distribution, waste generated in operations, business travel, employee commuting, leased assets, processing of sold products, use of sold products, end of life treatment of sold products, franchises, and investments. State "YES" if the company explicitly explains which categories it included and which categories it did not include, as well as why. State "NO" if you only find information about some categories included, but no additional information on the reasons why the other categories have been excluded. Include information on the challenges and possible solutions that companies apply to improve the categories coverage in the future.
Does the company report its annual progress of reducing GHG emissions to achieve its emission reduction or net zero targets?	Do not summarize the general climate and emission targets of the company. Focus on the precise information about the progress of its quantified GHG emission reductions, and whether this progress is in line with its general climate, emission or net zero target. If you only find information about the general climate targets of the company, answer "NO". If you find the precise information about the emission reductions so far, and whether this is in line with how the company expects to achieve its climate, emission or net zero targets, answer "YES".
Does the company report its absolute scope 1 GHG emissions for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 1 GHG emissions for the past 5 years. Answer "NO" if you do not find quantitative information or if the information does not cover the full 5 year period.
Does the company report its absolute scope 2 GHG emissions for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 2 GHG emissions for the past 5 years. Answer "NO" if you do not find quantitative information or if the information does not cover the full 5 year period.
Does the company report its absolute scope 3 GHG emissions for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 3 GHG emissions for the past 5 years. Answer "NO" if you do not find quantitative information or if the information does not cover the full 5 year period.
Does the company report a decline in its scope 1 GHG emission intensity for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 1 GHG emission intensity declining for the past 5 years. Answer "NO" if you do not find quantitative information, if the information does not deal with emission intensity, or if the information does not cover the full 5 year period.

Question	Expert-Centric Question Extension
Tracking	
Does the company report a decline in its scope 2 GHG intensity for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 2 GHG emission intensity declining for the past 5 years. Answer "NO" if you do not find quantitative information, if the information does not deal with emission intensity, or if the information does not cover the full 5 year period.
Does the company report a decline in its scope 3 GHG intensity for the past 5 years?	Acknowledge the fact that this information could also be conveyed in graphs and tables which you are not able to fully analyse. Answer "YES" if you find the precise quantitative information about the company's scope 3 GHG emission intensity declining for the past 5 years. Answer "NO" if you do not find quantitative information, if the information does not deal with emission intensity, or if the information does not cover the full 5 year period.
Does the company report the specific drivers and reasons for the company's observed actual GHG emission changes?	Do not include information about general climate or emission targets and activities of the company. Focus on the specific reasons spelled out by the company to explain why there has been an increase or decrease in emissions or emission intensity in the past until recently. State "NO" if you only find information about general emission reductions without linking them precisely with specific information about the relative importance to specific drivers of the change. State "YES" if such drivers and reasons for the emission change are explained, and if the information contains the quantified relative share of the drivers that is responsible for the decline or increase in the emissions.
Does the company report annual progress against its deforestation targets?	No additional guidelines
Does the company report the amount of climate aligned capex that supports its net zero transition?	No additional guidelines
Does the company report the amount of climate transition misaligned capex?	No additional guidelines
Does the company report the amount of climate transition aligned revenues that support the global net zero transition?	If you find information about the revenues from specific climate-related business branches of the company, summarize this information, but state "NO" if you do not find information about the full scale of the revenues that are climate transition aligned. State "YES" if you find specific information about the share of climate transition aligned revenues of the company in the recent past.
Does the company report the amount of climate transition misaligned revenues?	No additional guidelines
Does the company assess and report the alignment of its transition plan with its policy positions and its trade association's policy positions and lobbying?	No additional guidelines
Does the company report its engagement activities with the companies it invests in its own financial portfolio (including voting and proxy voting) undertaken in the relevant reporting period?	No additional guidelines

Table S.9: Expert-centric question extensions for Tracking questions

S.7 Human Evaluation

The human evaluation is structured to obtain insights into system quality, trustworthiness, and tool usage. This evaluation aims to collect qualitative, expert-based feedback. Thus, it does not represent an objective assessment, but it should inform users about expectations about the tool, applications, and potential improvement areas.

We designed two sets of questions. The first set of questions addresses the tool's performance on individual indicators. The tool processes items, i.e. question-answer-pairs corresponding to each indicator. As displayed in Table S.10, we ask for a range of questions about the quality of the answers provided by the tool for individual indicators (Q1-Q9). Furthermore, we ask the participants to assess the general impression and tool performance (Q10-Q15) (see Table S.11). Participants were asked to choose one option amongst the available ansers, i.e. multiple answers were not allowed. These questions can be mapped into the system quality, trustworthiness, and usage dimension as displayed in Table S.10.

As described in the text, for the pilot study we invite participants from 26 organizations spanning regulators/supervisors, banks, industry associations, NGOs, exchanges, academics, and investors. The participants stem from all over the world with a dominance in Europe. See Figure 2 for more details.

We ask every participant to submit at least three self-selected reports to the pilot study. Participants are also allowed to submit more than three reports. This aims to ensure that the experts also have prior interest and experience in analyzing the underlying companies. As a result, we obtain 93 reports from our participants. Figure S.7 shows the details of sectors and regions of the companies under investigation as well as the types and years of the report.

After submission, we analyze every report with our NLP tool. Then, we randomly assign the participants two indicators to assess from the sample of six most important indicators (see main text). Thus, the participants analyze at least two indicators for each of the three reports. Participants are allowed to assess more than the assigned indicator.

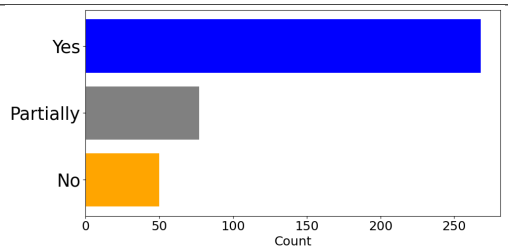
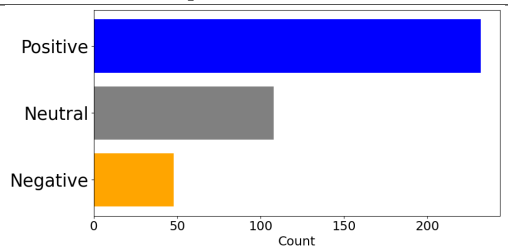
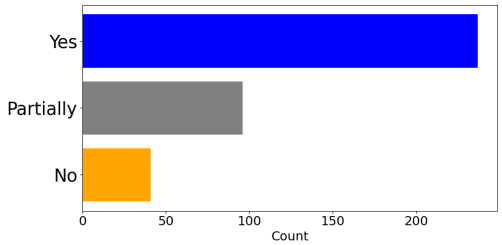
As a result of this process, we obtain 26 tool assessments and 396 indicator assessments. Table S.12 and S.13 show the detailed results. Overall, the tool reaches a very satisfactory performance. However, the participants also outline future improvement potentials.

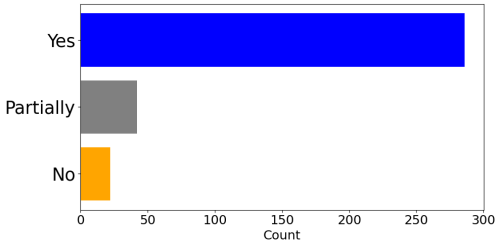
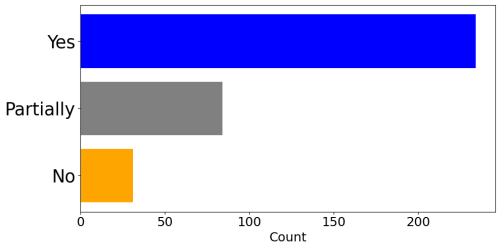
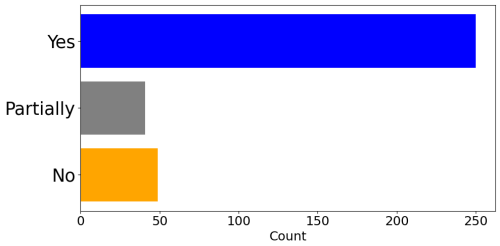
Question	Answer Options	Quality Dimension
Q1: If you had to answer the question, would you know what information to look for in a report?	- Yes - No - Partially	Usage
Q2: What is your first impression of the answer?	- Positive - Negative - Neutral	Usage, Trustworthiness
Q3: Does the model only summarize relevant content from the report to answer the question?	- Yes - No - Partially	System Quality
Q4: Does the model summarize relevant content correctly without making up information not contained in the report?	- Yes - No - Partially	System Quality
Q5: Do the sources provided support your trust in the model?	- Yes - No - Partially	Trustworthiness
Q6: Does the model cite all the information summarised from the report in the answer?	- Yes - No - Partially	System Quality
Q7: Does the model only cite pages from the report, which it uses to answer the question?	- Yes - No - Partially	System Quality
Q8: Does the model cite all the pages you think are most important to answer this question?	- Yes - No - Partially	System Quality
Q9: Does the model cite pages, which do not exist at all in the report?	- Yes - No - Partially	System Quality

Table S.10: Indicator assessment for the NLP tool answers by individual indicator.

Question	Answer Options	Quality Dimension
Q10: How do you like the tool in general?	- Very much - Somewhat - Not at all	Usage, Trustworthiness
Q11: The model provided answers which were mostly comparable to answers by...	- A person unfamiliar with the topic - A person who received a basic introduction to the topic - A person who has been working on the topic for 1-2 years - An expert in the topic with more than 2 years of analysis experience - An expert in the topic with more than 2 years of analysis experience who read the report carefully - An expert in the topic with more than 2 years of analysis experience who did not have enough time to properly read the details of the report	System Quality, Trustworthiness
Q12: I would like to use the model to...	- Be plugged into a possible risk-return analysis for an investment decision - Confirm my overall impression of how the company addresses the climate transition - Provide information on how the company addresses the climate transition - Provide the relevant information to assess the company's forward-looking business risks and opportunities - Provide the relevant information to assess the company's forward-looking business risks - Understand detailed aspects and nuances of the company's climate transition plan - Understand the broad picture of a company's transition strategy	Usage
Q13: The model answers are useful to...	- Be plugged into a possible risk-return analysis for an investment decision - Confirm my overall impression of how the company addresses the climate transition - Provide information on how the company addresses the climate transition - Provide the relevant information to assess the company's forward-looking business risks and opportunities - Provide the relevant information to assess the company's forward-looking business risks - Understand detailed aspects and nuances of the company's climate transition plan - Understand the broad picture of a company's transition strategy	Usage
Q14: The model needs to improve in the following ways to support my use case...	- Higher correctness of the answers - Less details provided in the answers - More details provided in the answers - Less irrelevant information in the answers - Always quoting the cited pages directly in the answer - Less missing out of relevant information in the report	System Quality, Usage
Q15: If I used the model, I would trust the information provided	- Yes - No - Partially	Trustworthiness
Q16: If I used the model, I would use the information provided and add the following quality checks:	(Free text only)	System Quality, Usage

Table S.11: Tool assessment for the NLP tool in general.

Question	Answers
Q1: If you had to answer the question, would you know what information to look for in a report?	 **Summary of respondent's criticism:** Most responses indicate a clear understanding of specific aspects to search for, such as ""CapEx alignment,"" ""external assurance,"" ""net zero policies,"" and ""GHG emissions reduction targets."" These responses suggest that certain participants were well-oriented towards identifying key elements of credible transition plans within reports. However, others expressed confusion or a lack of familiarity with specific terms or concepts like ""CapEx alignment"" and the general topic of transition plans, indicating a potential gap in understanding or knowledge. This variation suggests that while the tool might be helpful for those with some background or familiarity with the subject matter, it could be less accessible or intuitive for individuals lacking prior knowledge or expertise in climate transition strategies and reporting. Overall, the feedback points to a need for clearer guidance or educational components within the tool to accommodate users with varying levels of expertise and to ensure a broader understanding of how to identify and interpret the relevant information in reports.
Q2: What is your first impression of the answer?	 **Summary of respondent's criticism:** Several participants praised the tool for its accuracy and comprehensiveness, highlighting specific instances where the tool effectively connected complex pieces of information or addressed nuances in the transition plans. For example, one participant appreciated the tool's ability to connect the supervisory board's role with reduction goals as part of a transition plan, despite certain terms not being explicitly mentioned on the page. However, there were also criticisms, particularly regarding instances where the tool's answers were deemed misleading, incomplete, or failed to capture essential details. Some participants noted the tool's inability to correctly identify the presence or absence of specific elements within a company's transition plan, such as missing targets, baseline years, or the scope of GHG emissions covered. These criticisms suggest that while the tool can offer valuable insights, there is room for improvement in ensuring the accuracy and completeness of its analysis, especially in handling complex or nuanced information.
Q3: Does the model only summarize relevant content from the report to answer the question?	 **Summary of respondent's criticism:** Several participants appreciated the model's precision and conciseness in extracting relevant information, particularly in identifying specific goals, emissions data, or strategic plans directly related to the query. However, criticisms were common regarding the inclusion of irrelevant details, such as broader sustainability topics not directly answering the question, or speculative comments perceived as interpretive rather than strictly summarizing report content. Some responses highlighted the model's inconsistency, either overemphasizing certain aspects or missing crucial information, leading to partial relevancy in its summaries. This suggests a need for the tool to refine its focus on directly relevant content and reduce interpretative or supplementary information unless it directly supports the question. Overall, while the tool demonstrates capability in content summarization, there is room for improvement in ensuring relevance and limiting interpretive additions.

Question	Answers
Q4: Does the model summarize relevant content correctly without making up information not contained in the report?	 **Summary of respondent's criticism:** The participants' feedback on the AI tool's performance in summarizing relevant content without fabricating information not present in the reports shows a predominantly positive assessment, with a significant majority affirming its accuracy. However, notable exceptions were highlighted where the tool either partially met the criteria or failed. These exceptions often related to the tool's interpretation or addition of judgments not directly derived from the source material. For instance, issues were raised concerning the tool's misinterpretation or misleading representation of certain targets or achievements, such as incorrectly summarizing emission reduction targets or mischaracterizing the scope of certain corporate actions. Some participants also pointed out that the tool might go beyond summarizing to making unwarranted assessments, potentially veering into subjective analysis not supported by the report contents. This suggests that while the tool is largely effective in capturing and summarizing report contents accurately, there is room for improvement in ensuring that all interpretations and judgments are well-grounded in the source documents.
Q5: Do the sources provided support your trust in the model?	 **Summary of respondent's criticism:** Several participants highlighted specific issues with source relevance, citing examples where sources were either not provided, irrelevant, or lacked comprehensive coverage on the topic addressed by the question. Notably, some responses pointed out missed sources that could have substantiated the model's answers better, suggesting a need for the tool to encompass a wider range of relevant data. Concerns about the completeness and accuracy of the information underline a critical view on the trustworthiness of the model, with participants suggesting that while the model might identify correct pages or concepts, it often overlooks detailed verification or inclusion of all pertinent information. This feedback suggests that for the tool to enhance user trust, it must not only identify relevant sources but also ensure comprehensive coverage and contextual relevance, providing a more detailed justification for its answers.
Q6: Does the model cite all the information summarised from the report in the answer?	 **Summary of respondent's criticism:** Several responses indicate that the tool either missed citing some sources altogether or provided citations that didn't directly support the summary provided. Key observations include a lack of direct mention of sources, missing relevant page numbers, and incorrect or irrelevant citations. Some participants noted that while the tool mentioned sources, it sometimes omitted significant details or relevant pages that contained crucial information. There were also instances where the tool partially cited sources but failed to provide a comprehensive view by excluding important references. This suggests that while the tool might capture some relevant information, there's room for improvement in ensuring completeness and accuracy of citations, highlighting the importance of direct, precise, and inclusive referencing in summarizing and analyzing reports.

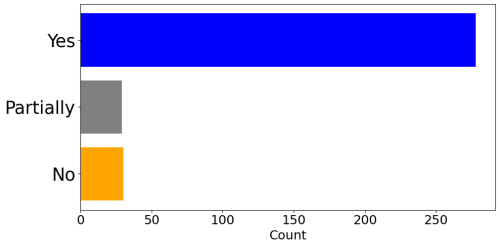
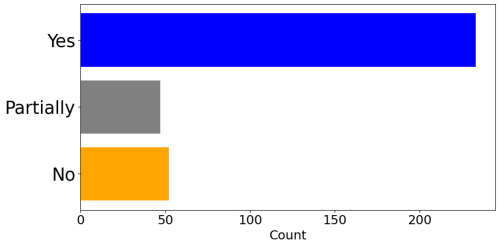
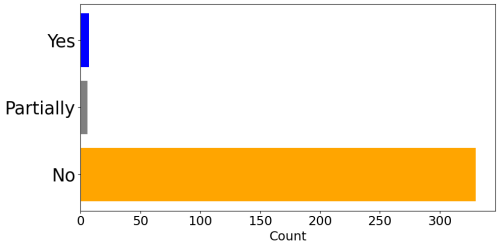
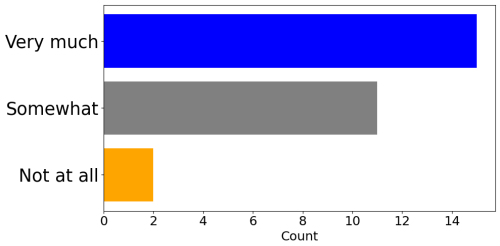
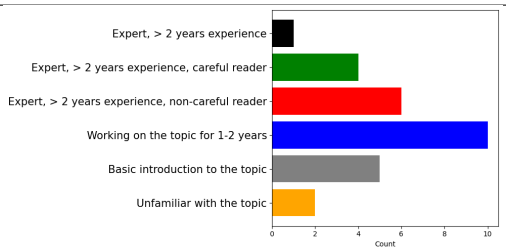
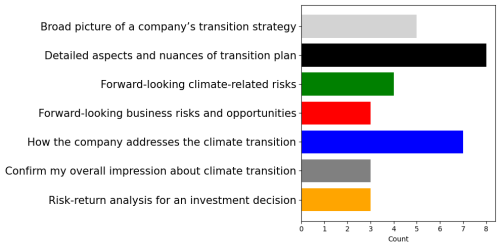
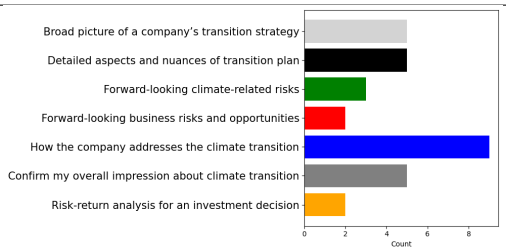
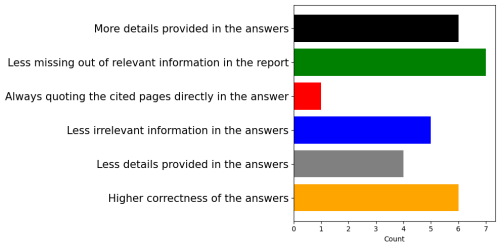
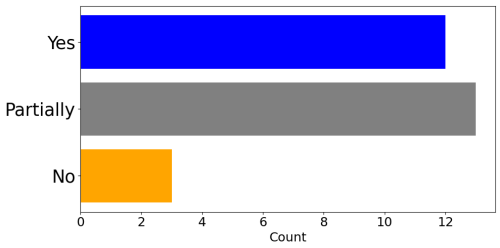
Question	Answers
Q7: Does the model only cite pages from the report, which it uses to answer the question?	 **Summary of respondent's criticism:** Several respondents indicated that the model either cited irrelevant pages or missed citing critical pages that were directly relevant to the question. Some specific feedback pointed out that certain pages mentioned by the model had no relevance to the query at hand, suggesting a mismatch between the model's citations and the actual content required to answer the question. Others noted that while the model cited pages, it either included unnecessary ones or omitted crucial information, indicating a gap in the model's understanding or retrieval capabilities. A few responses were more nuanced, acknowledging partial success but highlighting inconsistencies or the inclusion of seemingly irrelevant pages. Overall, this feedback underscores a need for improvement in the model's accuracy and relevance in citing pages from reports, ensuring it focuses only on pertinent information to address the query effectively.
Q8: Does the model cite all the pages you think are most important to answer this question?	 **Summary of respondent's criticism:** Several participants noted missing pages that contained relevant information, suggesting that the model might overlook critical content or fail to comprehensively assess the report's entirety. Concerns about missing pages indicate a need for the model to broaden its search and citation strategy to ensure it captures all pertinent information. Some feedback pointed to alternative pages that could provide more detailed or accurate information to support the model's answers better. This diversity in feedback underscores the complexity of identifying and citing the most relevant pages in extensive reports and the importance of continually refining the model's algorithms to improve its accuracy and relevance in citations.
Q9: Does the model cite pages, which do not exist at all in the report?	 **Summary of respondent's criticism:** Among the few who provided an explanation, key insights emerge. For instance, one answer highlights a lack of source material as a reason for their response, suggesting some participants might have found the assessment challenging due to insufficient information. Two answers specifically mention issues related to page citations, noting a repetitive mention of a page in a given cell, or stating no pages were cited. This indicates a perceived discrepancy in the model's citation practice or an error in the report's page numbering. Two remarks about unclear references to pages suggest confusion about whether the citations referred to physical or electronic document formats. Overall, these explanations point to a need for clearer guidelines on citations and perhaps better access to or organization of source materials for more accurate assessment by participants.

Table S.12: Item assessment result for the NLP tool

Question	Answers
Q10: How do you like the tool in general?	 **Free-text comments:** No free-text comments allowed.
Q11: The model provided answers which were mostly comparable to answers by...	 **Summary of respondent's criticism:** The free-text explanations suggest a generally positive, though nuanced, perception. Several respondents highlighted the tool's ability to grasp key information and offer basic analysis on climate transition plans, indicating its utility for individuals with a basic to intermediate understanding of the topic. However, criticisms emerged regarding the tool's depth of analysis, tone, relevance of commentary, and its tendency towards caution, suggesting room for improvement in delivering more expert, assertive, and concise analyses. The tool was recognized for its informed and authoritative voice, yet some felt it could benefit from a clearer stance on assessments and a refinement in distinguishing between different types of reports. This feedback points towards the tool's potential as a helpful resource, especially with enhancements in precision, expertise-level analysis, and succinctness in responses.
Q12: I would like to use the model to...	 **Summary of respondent's criticism:** Participants generally seek deeper insights into companies' climate transition plans, including the ability to understand nuances, access precise data, and identify discrepancies across reports. There's also an interest in the tool's potential for risk-return analysis in investment decisions and the evaluation of forward-looking climate-related and business risks. Notably, some participants express a desire for the tool to aid in broader strategic understanding, such as comparing company reports, flagging inconsistencies, and assessing the public perception of their initiatives. The feedback suggests that while the tool is seen as a potentially valuable asset for quick assessments and enhancing work processes, users also crave comprehensive analyses to feel confident in their decisions or presentations. This indicates an appetite for a tool that not only provides summaries but also deep, actionable insights.

Question	Answers
Q13: The model answers are useful to...	 **Free-text comments:** Few participants offered specific insights that reveal areas for potential improvement and trust concerns. One response highlighted the need for the tool to address climate transition credibility more explicitly, suggesting a desire for a deeper analysis of how companies' actions align with their stated commitments. Another response pointed to a need for gap analysis towards identifying greenwashing, indicating an interest in more nuanced evaluations of companies' sustainability claims. A critical note was also made about the current trustworthiness of the tool's answers, with one participant stating they would still feel the need to verify disclosures independently. Lastly, there was a constructive suggestion about using the tool to analyze how external parties view a company's publicized sustainability efforts, which could help in refining and improving communication strategies.
Q14: The model needs to improve in the following ways to support my use case...	 **Summary of respondent's criticism:** Participants noted issues with the tool either omitting crucial information or incorporating irrelevant details, particularly in regards to interpreting data in tables and understanding implications of financial allocations. There's a desire for clearer delineation between direct quotations, summaries, and analytical judgments, as well as for more targeted questioning to reduce the risk of "greenwashing" by accepting weak disclosures as sufficient. Suggestions for enhancements include more granular questioning, breaking down information for ease of understanding, and ensuring that answers directly cite sources to facilitate verification.
Q15: If I used the model, I would trust the information provided	 **Summary of respondent's criticism:** Concerns were raised about the tool's ability to handle complex, sector-specific issues and in-depth qualitative analysis, suggesting it performs better with general information and broader overviews. Some responses highlighted inaccuracies in model interpretations and a desire for a more comprehensive assessment beyond climate and environmental aspects. Additionally, the trustworthiness of the tool's outputs seems to be contingent on their alignment with established players in the field, like DJSI and CDP.

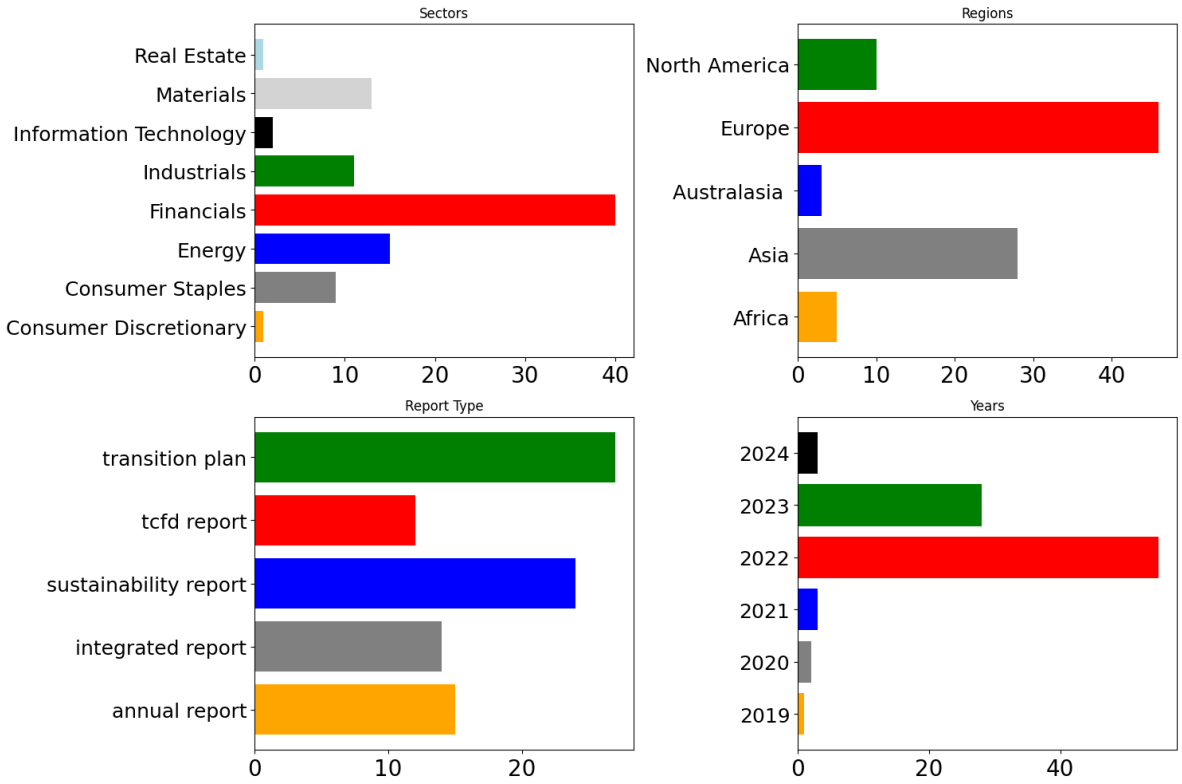


Figure S.7: Overview of the 93 reports in the evaluation.

Question	Answers
Q16: If I used the model, I would use the information provided and add the following quality checks:	**Free-text comments:** Participants highlighted a range of additional quality checks they would implement to enhance the trustworthiness and utility of the AI tool's outputs. Common themes include the necessity for cross-checking information across companies, sectors, and countries to ensure fairness and consistency in analysis. There's a clear demand for the model to accurately identify and compare specific elements, such as carbon offsetting practices between different companies. Users expressed the need for detailed verification of negative responses and alternatives considered, suggesting a deeper dive into qualitative analysis and materiality specific to each industry. Moreover, the feedback indicates a desire for more accessible, digestible summaries and targeted information to address specific inquiries. Participants also suggested integrating checks on the accuracy of numerical data and the relevance of cited sources, alongside cross-referencing indicators for a comprehensive assessment of transition plans' credibility. This feedback underscores a cautious approach, valuing accuracy, comprehensiveness, and user-friendliness in the tool's application.

Table S.13: Tool assessment results for the NLP tool.

S.8 Word clouds

To provide additional insights into the main topics covered in the most and least disclosed indicators, we display word clouds based on the frequency of terms used in the corresponding questions. Figure S.8 confirms our interpretation of the factors driving the divergence in performance between indicators. While companies are more likely to report a target for emission reduction, they tend to not disclose information related to offsets and fossil fuels.

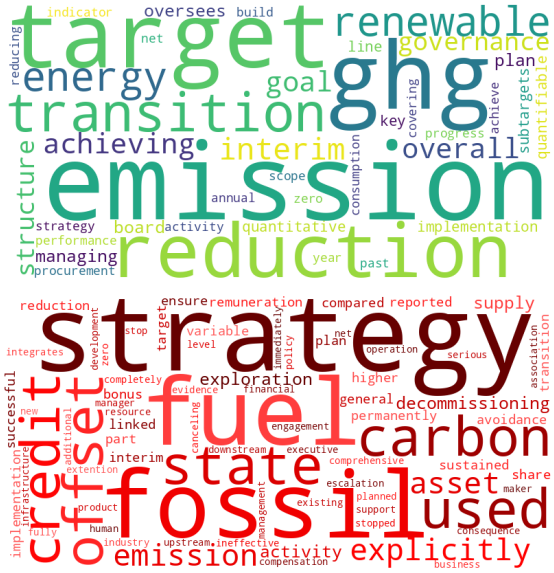


Figure S.8: Word clouds for the 10% best (green) and worst (red) performing questions.

Net Zero Transition Plans: Red Flag Indicators to Assess Inconsistencies and Greenwashing

Tobias Schimanski, Julia Bingler, Chiara Colesanti Senni, Jingwei Ni, Markus Leippold

University of Zurich, ETH Zurich, University of Oxford

Problem Set and Solution Approach

Current Set of Problems

- Companies need to transition towards a low carbon economy
- Vast amounts of recommendations, principles, guidelines for disclosure
- Companies disclose, but no one exactly knows how good or bad

Vast amounts of data, information asymmetry

Solution Approach

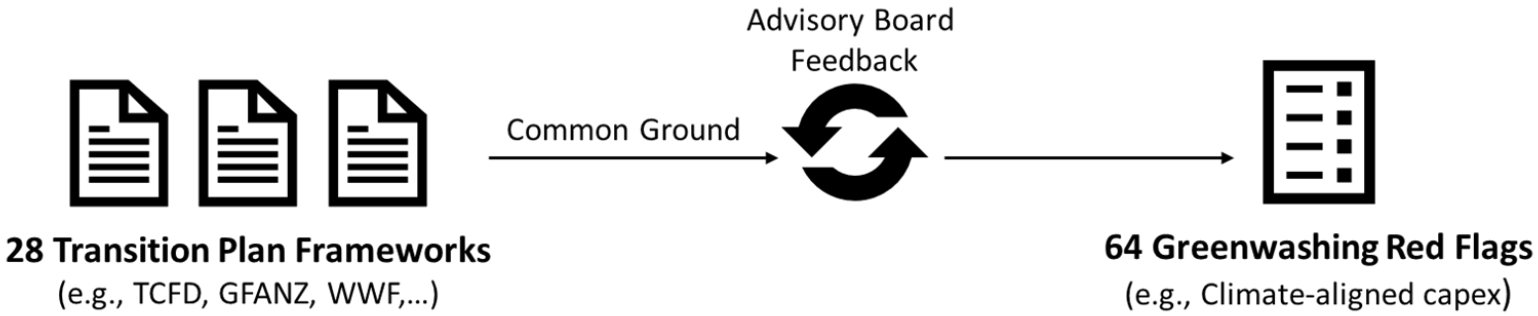
1. Create common ground between disclosure frameworks
2. Create and validate AI tool that assesses company disclosures
3. Apply AI tool on company disclosures to assess patterns

Automatic approach of assessing companies on scale

1. Assessment Indicators

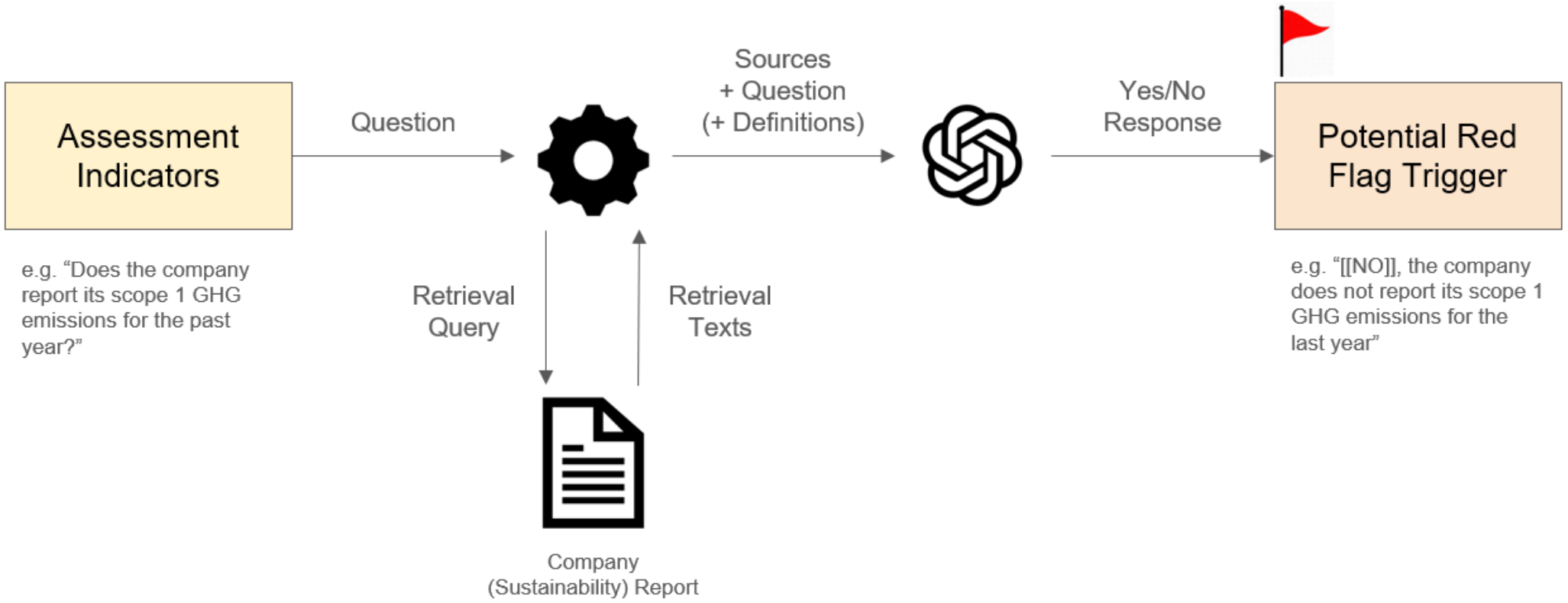
Identify indicators to assess transition disclosures based on

- External consistency: ambition and feasibility of transition plans
- Internal consistency: credibility of transition plans toward achieving a net-zero business strategy



Red flags to signal that transition plans underperform against some of the selected criteria and the risk of greenwashing

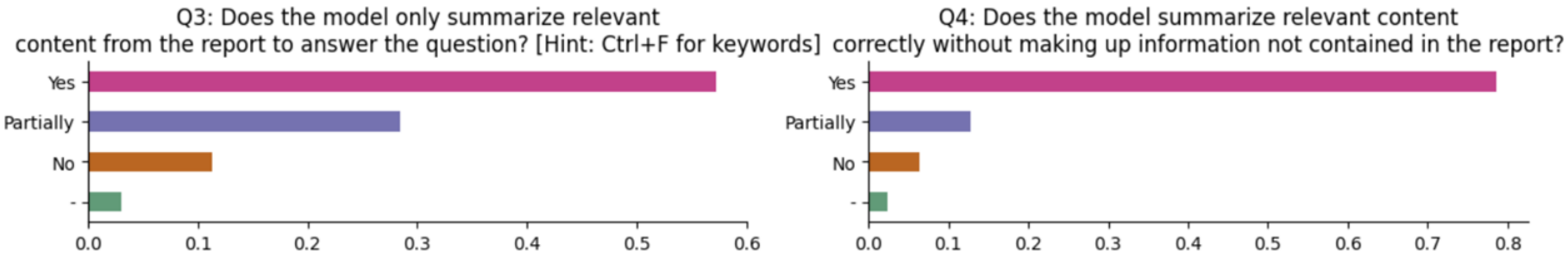
2. Retrieval Augmented Generation



Focus on traceability, verifiability and transparency

2. Expert-Centric Validation Approach

Pilot Study: 26 institutions spanning central banks, NGOs, investors, exchanges, regulators/supervisors, etc.: Quantitative and Qualitative Validation

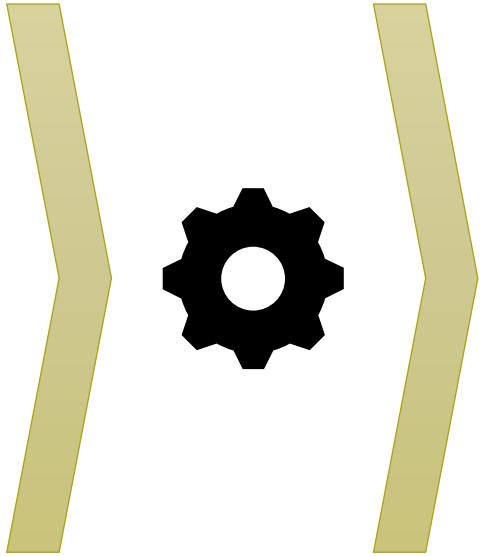


Key Insights:

- Very Satisfactory level of system quality: source quality, citing, very limited hallucination
- Trustworthiness of the system depends on external verifiability and user needs
- Usage potentials are vast, improvement through sector- or investor-specialisations of the tool

3. CA100+ Study

Companies
Analyse reports of the highest emitters in the world



Key Insight

- Indicators relating to “talk” are most often fulfilled (commitments, plans)
- Indicators related to “walk” are least often fulfilled (achievements, details on execution)

=> Huge amount of reporting but selective communication (?)

Outlook and Limitations

Data

- Reliance on “solely” sustainability reports
- Availability of reports

Model and Indicators

- Technical improvements will enable huge advancements
- Flexibility of the Indicators

Greenwashing

- Indicators show self-reported perspective and deliver “indication”
- Third-party and human verification

Thank you for your attention!

Working Paper to be out soon

Feel free to follow on LinkedIn

Tobias Schimanski