Structuring Sustainability Reports for Environmental Standards with LLMs guided by Ontology

Aida Usmanova
Leuphana University of Lüneburg
aida.usmanova@stud.leuphana.de

Ricardo Usbeck
Leuphana University of Lüneburg
ricardo.usbeck@leuphana.de

Abstract

Following the introduction of the European Sustainability Reporting Standard (ESRS), companies will have to adapt to a new policy and provide mandatory sustainability reports. However, implementing such reports entails a challenge, such as the comprehension of a large number of textual information from various sources. This task can be accelerated by employing Large Language Models (LLMs) and ontologies to effectively model the domain knowledge.

In this study, we extended an existing ontology to model ESRS Topical Standard for disclosure. The developed ontology would enable automated reasoning over the data and assist in constructing Knowledge Graphs (KGs). Moreover, the proposed ontology extension would also help to identify gaps in companies' sustainability reports with regard to the ESRS requirements. Additionally, we extracted knowledge from corporate sustainability reports via LLMs guided with a proposed ontology and developed their KG representation.

1 Introduction

Presently, environmental and social justice are the main challenges demanding our attention, with a focus on transparency. As a result, there is an increasing demand for organizations to disclose non-financial information, particularly in sustainability reports. To regulate the disclosed information, companies can adhere to Environmental, Social and Governance (ESG) regulations. While there exist multiple ESG frameworks, such as the widely adopted Global Reporting Initiative (GRI) and Sustainability Accounting Standards Board (SASB), effectively addressing all policy requirements and keeping track of all standards presents a challenging task. The most recent standard, released by the European Union Corporate Sustainability Reporting Directive (CSRD), is the European Sustainability Reporting Standard (ESRS). ESRS aims to disclose the company's strategy to mitigate negative impact and to align with the Paris Agreement. Incorporating ESRS standards in annual reports will become mandatory in the European Union, starting from 2025.

The transition from well-known frameworks like GRI to ESRS involves several modifications. While the majority of GRI disclosures have a corresponding ESRS equivalent, the latter demands more data points and detailed information. Specifically, financial materiality analysis in terms of risks and opportunities the company's environmental impact entails.

Advancements in Large Language Models (LLMs) enable fast and effective processing and extraction of relevant information from the textual sources (Brown et al., 2020). Coupling LLMs with Knowledge Graphs (KG) allows us to conveniently represent unstructured reports in a structured format automatically. Since reading and analysing non-financial disclosure reports can become a long and cumbersome task, leveraging Deep Learning models provides a means for fast automatic analysis of large numbers of information. Making it feasible to transition the existing GRI and ESG reports into ESRS format, as well as pinpoint requirements that need to be examined and addressed in a more detailed fashion.

The primary objective of this study is to automate the extraction of climate change-relevant information disclosed by companies in their non-financial reports and investigate the seamless transition to ESRS reporting standards. Large unstructured disclosure reports are transformed into structured graph-based representations for further analysis. The proposed approach is based on an extension of the Text2KGBenchmark (Mihindukulasooriya et al., 2023) Knowledge Graph generation from text guided by an ontology. Therefore, we extended the existing OntoSustain (Zhou and Perzylo, 2023) ontology to include the ESRS Topical Environmental Standards. We evaluate our approach through human annotation and see satisfying results. A link to the source code will be added after the review phase.

2 Related Work

Developing a non-financial disclosure report analysis involves collecting information from multiple textual sources, which may lead to confusion due to the number of reporting requirements. Ontologies are capable of modelling complex domain knowledge and mitigating natural language ambiguities (Navigli et al., 2003), hence, they are perfectly suitable for the task. Ontologies are defined as "a means to formally model the structure of a system, i.e., the relevant entities and relations that emerge from its observation, and which are useful to our purposes" (Guarino et al., 2009). (Zhou and Perzylo, 2023) developed an ontology - OntoSustain - that models sustainability domain knowledge and offers a platform for companies data collection process.

Knowledge Graphs (KGs) offer a structured database for storing and representing the information from multiple sources, which facilitates a wide range of tasks, like semantic search, explainable AI, question answering, information retrieval (Hogan et al., 2021). KGs contain the instance data for information modelled according to a specific ontology.

LLMs trained on large corpora reach state-of-the-art performance across multiple NLP tasks due to prompt engineering (Brown et al., 2020). LLMs' information extraction capabilities have shown to provide relevant structured information for KG construction (Carta et al., 2023; Pan et al., 2023; Zhu et al., 2023; Trajanoska et al., 2023; Meyer et al., 2023). The resulting KG represents extracted information as concepts and the relations between them as edges (Reinanda et al., 2020).

(Bronzini et al., 2023) proposed an approach for extracting structural insights related to ESG aspects from sustainability reports by leveraging LLMs, In-Context Learning and Retrieval Augmented Generation (RAG). The statistical analysis proved disclosure similarities between companies within a sector and region. The study also analysed the impact of ESG ratings on companies. Unlike (Bronzini et al., 2023) study, our focus is on structuring non-quantifiable report aspects.

To the best of our knowledge, to date, there have not been any studies conducted to represent and analyse ESRS standards for disclosure reports, particularly ESRS 2 Topical Standard, as KGs. Hence, our work aims to address this challenge and offers an ontology and a KG construction method for representing ESRS 2 Topical Standard.

3 Methods

This section outlines the ontology design, dataset selection and KG construction procedures.

3.1 Ontology extension

The OntoSustain ontology (Zhou and Perzylo, 2023) incorporates sustainability indicators from GRI and ESRS 1 reporting standards. The ontology covers the company's daily business activities, sustainability domain knowledge and the reported sustainability indicators. ESRS 1 general requirements are also included in the OntoSustain design.

In this study, we extended OntoSustain ontology to include ESRS 2 general disclosures, specifically Topical Environmental Standards. Topical Standard disclosure requirements are described by four reporting areas. Namely, Governance (GOV), Strategy (SBM), Impact, risks and opportunities (IRO), Metric and targets (MT). The Topical Environmental Standards are categorized into 5 topics: climate change (E1), pollution (E2), water and marine resources (E3), biodiversity and ecosystems (E4), and circular economy (E5). We tested the proposed ontology on the E1 category.

Figure 1 depicts an extended version of OntoSustain. The extensions are coloured in blue. The Topical Standard consists of 5 aspects that correspond to reporting areas, i.e., IRO, SBM, MT. Each aspect has a description and textual information extracted from the report. The following aspect definitions regarding climate change E1 category were considered in this study:

- Impact: Negative impact on climate change from a company's activities that the company addresses in the report.
- Risks: Material risks from impact on the climate change.
- Opportunities: Financial materiality from company's activities related to climate change.
- Strategy: Company's strategy and business model in line with the transition to a sustainable economy.

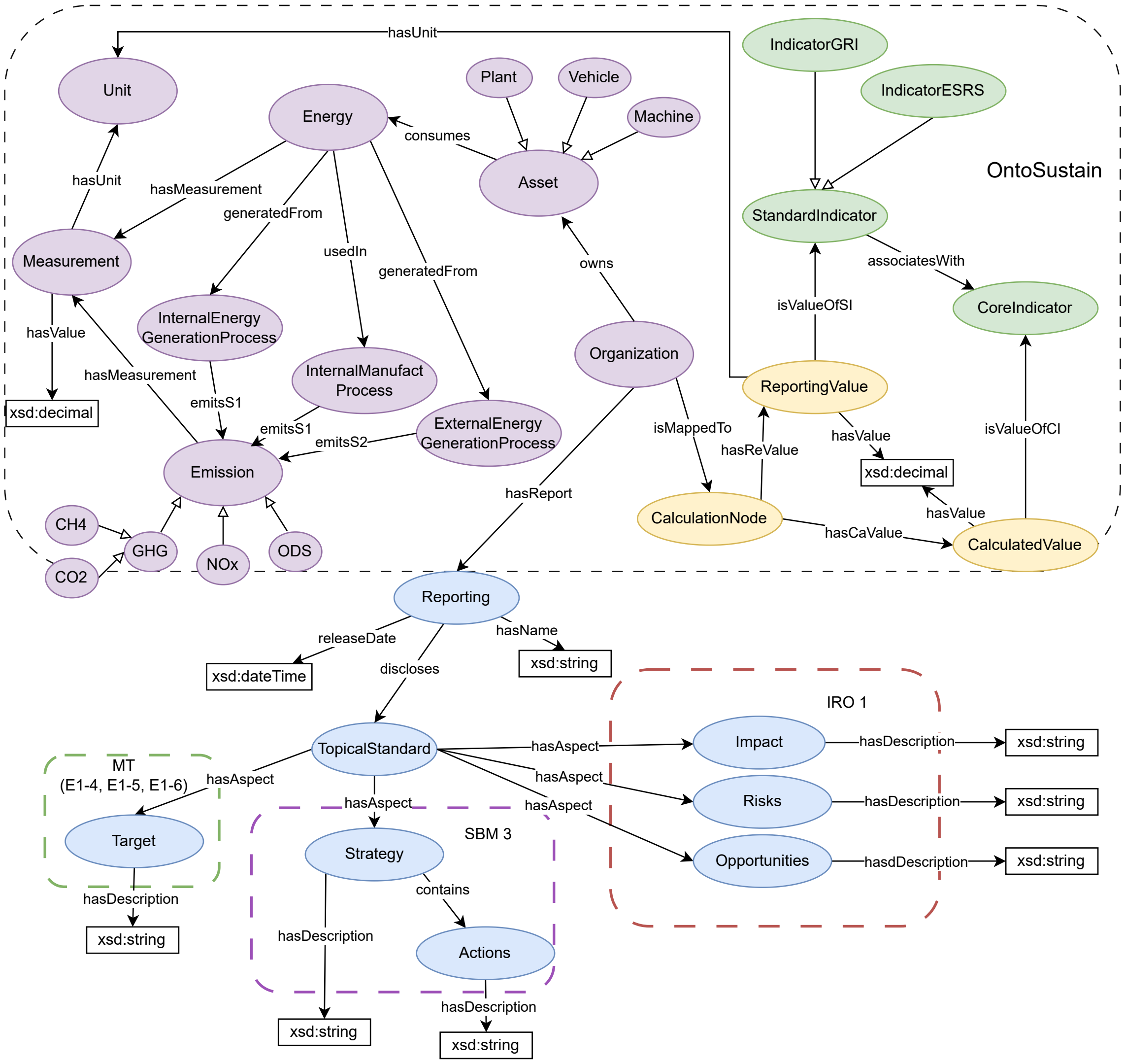

Figure 1: OntoSustain extended: added reporting areas for ESRS 2 Environmental Standards.

- Actions: Actions and resources in relation to material sustainability matters.
- Targets: Company's goals towards a sustainable economy.

3.2 Dataset

As mentioned earlier, (Bronzini et al., 2023) focused on extracting insights from non-financial disclosure reports and analysing their ESG standards' text. The authors generously provided access to 124 pre-processed sustainability reports. The majority of the reports were from North American companies. Given the focus of our study on the ESRS standard, we limited our dataset to EU-based companies, which resulted in 14 companies, see Table 1. The selected companies represent diverse industry sectors classified according to Global Industry Classification Standard (GICS). The majority of the reports were released for the 2021 fiscal year, with 3 reports being for 2020 and 1 for the 2017 fiscal year.

3.3 Knowledge Extraction and KG Construction

(Mihindukulasooriya et al., 2023) proposed an approach for KG construction guided by ontologies. The prompts for the LLMs are automatically constructed from ontology descriptions such as concepts, relations and domain constraints. Additionally, input and output examples are provided in the prompts to allow for few-shot extraction.

The (Mihindukulasooriya et al., 2023)-method involves several challenges such as automatic

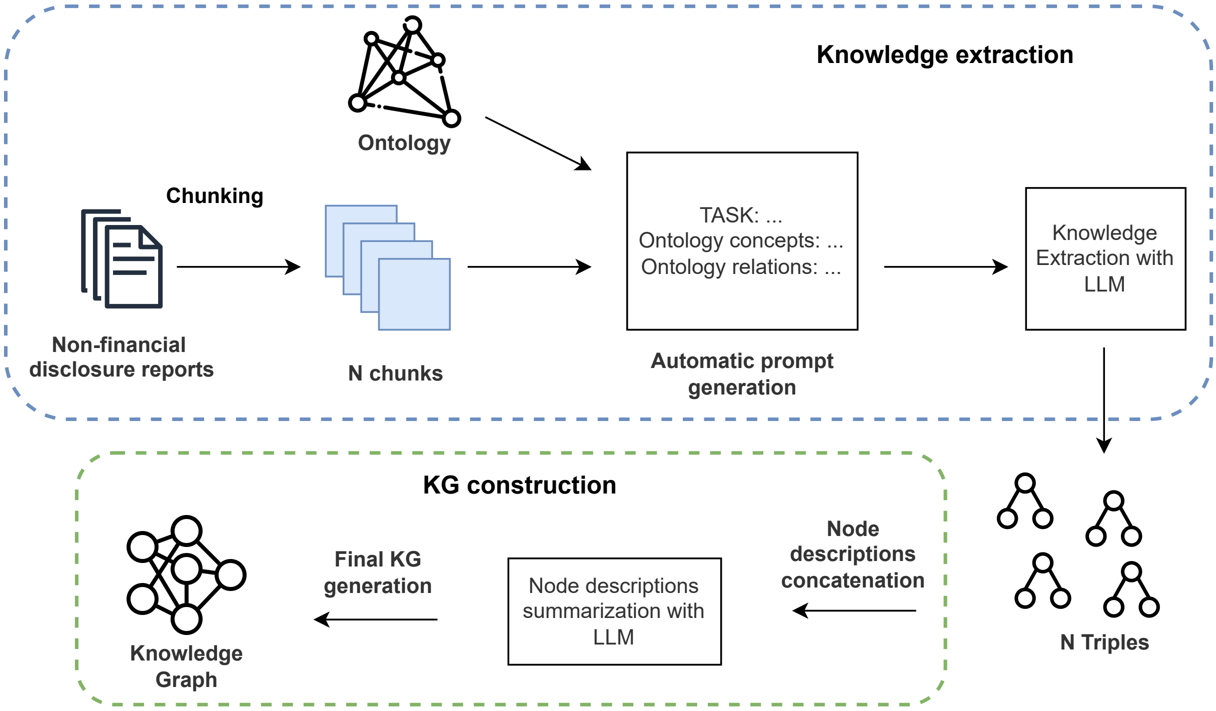

Figure 2: Study workflow

GICS Sector	Companies
Industrial	Airbus, Poste Italiane
Financials	Deutsche Bank, Santander Bank, Assicurazioni Generali
Communication	Telecom Italia
Healthcare	Bayer
Materials	ArcelorMittal, Lufthansa
Energy	Eni, Royal Dutch Shell, TotalEnergies
Utilities	Enel, Uniper

Table 1: 14 EU-based companies and the corresponding Global Industry Classification Standard (GICS) sectors.

prompt generation, relevant demonstration example selection, and addressing the LLM hallucination problem. That is, the generated prompt should present the ontology and its relations effectively while being efficient and descriptive enough. Providing helpful input and output examples can significantly improve the resulting output of the model. LLM should accurately extract relevant facts and not introduce new concepts and relations.

To ensure the quality of extracted information, we consulted with a macroeconomics expert researcher to improve the prompt with relevant examples by manual prompt engineering. As an example of input, we took a publicly available Siemens sustainability report for 2022. Together with the expert, we extracted all aspects relevant to climate change to give them as an example output. The prompt example for knowledge extraction is shown on Figure 5 in Appendix A section. The prompt provides a task description, followed by the relevant context. The context involves ontology concepts, ontology relations, example sentences and an example output.

Figure 2 demonstrates how the knowledge is extracted from the sustainability reports, guided by the extension of OntoSustain ontology. Given the length of corporate reports, which are large pieces of textual information, we utilised NLTKTextSplitter[1] for content-aware chunking. Such sentence tokenizer considers the content and the nature of the human language data, helping to split the text into sentences that are more meaningful chunks. The chunks and ontology then construct a prompt for an LLM input. Given report text is automatically divided by tokenizer into N chunks, depending on the size and the content of the document, which results in ontology concepts and relations being extracted N times (An extracted knowledge example is shown on Figure 7 in the Appendix A section). The Objects with a similar Subject and Relation are then concatenated into one text.

[1] https://api.python.langchain.com/en/latest/nltk/langchain_text_splitters.nltk.NLTKTextSplitter.html

Concatenated aspect descriptions (i.e. Impact, Risks, Opportunities, Strategy, Actions, Targets) and aspect definitions are then passed to the LLM for a summarization task. For Organization name, Reporting name and release date, we selected the most common value via majority voting.

As an LLM, we relied on GPT-4 (OpenAI, 2023) with the LangChain framework (Chase, 2022) for implementation. The experiments were done from March to April 2024.

4 Evaluation

Since there are no human-labelled or classifier-trained ESRS Topical Standard examples or training data, we asked three annotators to label the extracted Topical Standard aspects. Firstly, annotators were given a brief explanation of the conducted study and Topical Standard aspects definitions. Aspects extracted from the "Siemens sustainability report 2022" were given as examples. Later, annotators were asked to label 6 aspects from 14 reports, a total of 84 entries, based on how well the extracted descriptions match the definition of the respective aspect from the climate change Topical Standard. We also asked annotators to indicate if they think the aspect description is too general or vaguely written. Our annotation guide will be published after review, too.

All three annotators completed the informed consent procedure, demonstrating their understanding and willingness to participate in the study. Annotators provided scores individually and did not interact with each other. Figure 6 shows a screenshot of the spreadsheets form where annotators were asked to provide feedback, see Appendix A section.

Inter-Annotator Agreement (IAA) (Viera et al., 2005) assesses the level of agreement between multiple annotators in their evaluations of topical match. Specifically, we use Cohen's Kappa (κ) to quantify the degree of agreement among the annotators.

The formula for Cohen's Kappa is given by:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{1}$$

where:

$$P_o = \frac{\text{Number of agreements}}{\text{Total number of annotations}}$$

$$P_e = \sum_i \left(\frac{\text{Total annotations by annotator } i}{\text{Total number of annotations}} \right)^2$$

Aspect	A 1	A 2	A 3
Strategy	14	14	14
Impact	9	11	9
Risks	12	14	14
Opportunities	8	12	11
Actions	14	13	11
Targets	12	14	13

Table 2: Number of Topical Standard aspects classified as match by annotators.

Here, P_o represents the observed agreement, which is the proportion of times the annotators agree. While P_e represents the expected agreement, which is the hypothetical probability of agreement occurring by chance.

We calculated the Pairwise Cohen's Kappa, then averaged these values across all pairs of annotations to obtain an overall agreement measure. The value of κ ranges from -1 (complete disagreement) to 1 (complete agreement), with a value of 0 indicating agreement by chance. In our case, the average Cohen's Kappa score across all pairs of annotators is 0.512, which suggests a moderate level of agreement among the annotators. Annotation results per aspect and annotator are presented in Table 2.

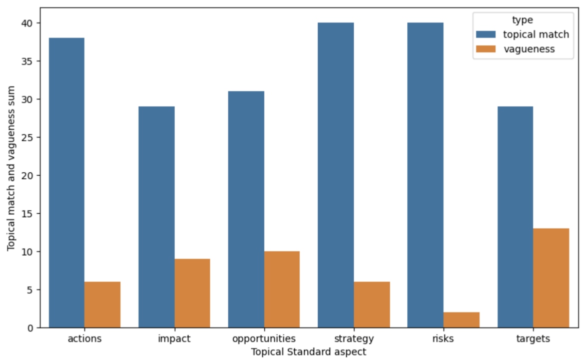

Figure 3: Topical match and vagueness sum per aspect.

5 Experimental Results

As mentioned above, the calculated Cohen's Kappa resulted in 0.512, a moderate agreement. Factors such as task complexity, ambiguity in annotation guidelines or differences in annotators' expertise could have influenced the outcome. The IIA results could potentially change when having more participants for annotation, however, our study had limited resources and missed more experts from the field of economics.

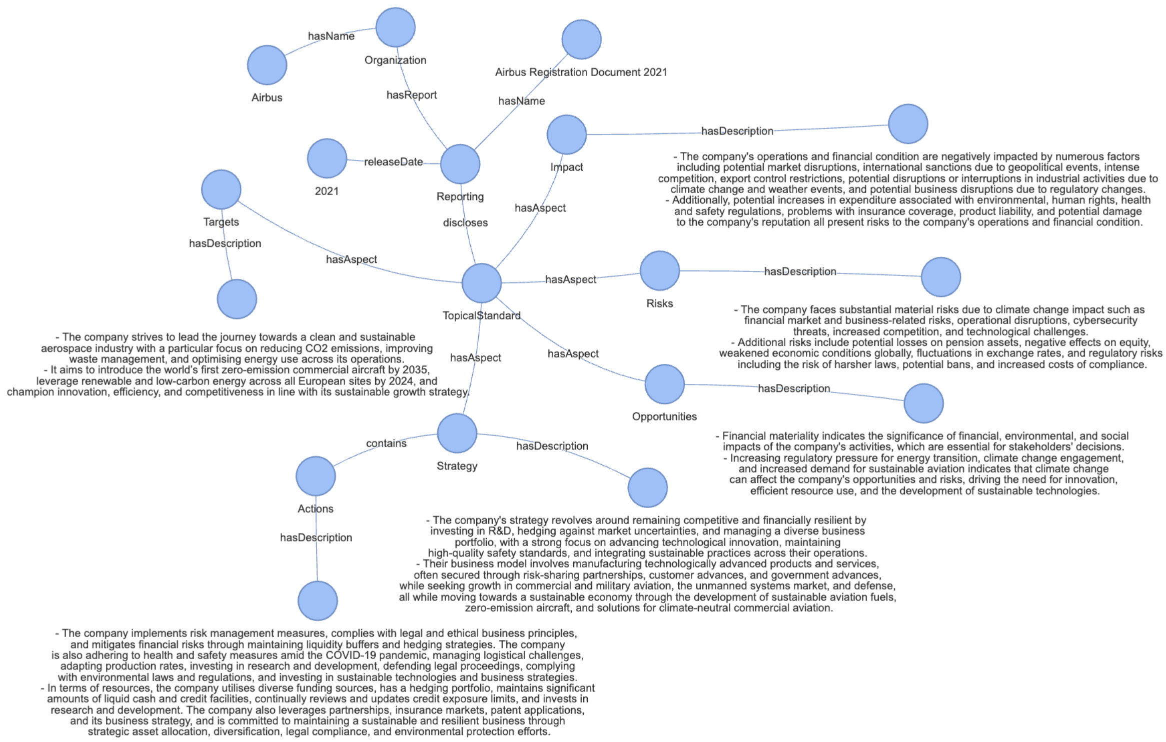

Figure 4: KG output example based on Airbus annual report.

The LLM was able to accurately extract organization name, reporting name and its release year for all 14 reports.

As seen from Table 2, *Strategy* aspect extracted from all reports have been classified as match by all annotators. *Risks*, *Actions* and *Targets* also have a high number of matches by all annotators. While *Impact* and *Opportunities* have the least number of matches according to annotators.

In most cases where *Impact* was classified as not match, the aspect summary was reporting either the impact of climate change on the organization, which is more related to *Risks* aspect, or actions the organization takes to be more sustainable, which belongs to the *Actions* aspect. *Opportunities* refers to significant financial impacts that climate change brings into the organization structure and economic decisions. Hence, many mismatched summaries were a mix of *Actions* and *Risks*.

Figure 3 depicts the number of topical match and vagueness results per aspect. Aspects that have the least topical match are also the ones containing the most vague descriptions. *Targets* were commonly labelled as vague due to the absence of quantifiable and measurable goals. Many *Impact* entries were considered vague because the descriptions included the impact of climate change on the organization, as well as the impact of the organization on climate change. On the contrary, *Opportunities* provided the definition of financial materiality in multiple entries, which was considered too general by annotators. Clearly, both aspects require a much more descriptive demonstration example during the Knowledge Extraction phase.

Overall, the suggested ontology extension and the ontology-guided prompting technique demonstrated reasonable results. An example KG extracted from the Airbus report is shown in Figure 4.

6 Conclusion

We present a novel method for extracting and structuring information from disclosure reports. Representing knowledge in such a way would offer more transparency and reusability of the data for further analysis. In this study, first, we designed an extension for existing ontology to represent ESRS 2 Topical Standard information. Second, we also tested the proposed extension by prompting GPT-4 with ontology guidance and expert prompt engineering. The results demonstrated that the proposed extension adapts well to many aspects, namely Strategy, Risks, Actions and Targets. While Impact and Opportunities require better description and representation.

Limitations

Processing documents is a very tedious task due to the size of textual information. Dividing such documents leads to a large number of chunks, with the longest document reaching 764 chunks in our case. Based on the selected Knowledge Extraction approach, each chunk needs to be passed together with the ontology separately. This leads to the knowledge extraction taking a long time to be executed and a huge invest for a small research lab. Hence, better pre-processing and chunking techniques are yet to be developed to use LLMs on corporate documents.

The extended ontology still requires extensive evaluation using data from more disclosure reports. The ontology also needs to be assessed with other Topical Standards from Environmental, Social and Governance sectors. The generated KGs need to be validated for consistency and completeness by the experts in the field. This suggests a human-in-the-loop approach, where a trained expert will customize a descriptive example for extraction or verify the validity of extracted data. This could be a potential development path for processing and analysing corporate reports.

Ethics Statement

Our research focuses on the application of LLMs, KGs and ontologies to assist companies in analysing sustainability reports. While the benefits of this technology are clear, it is essential to acknowledge and address potential ethical considerations.

Firstly, the reliance on LLMs to extract and interpret information from corporate sustainability reports may inadvertently perpetuate biases present in the source data. These biases can affect the fairness and accuracy of the generated knowledge graphs and automated reasoning outcomes, e.g., representing popular companies better than unknown companies.

Secondly, the automation of sustainability reporting carries the risk of over-reliance on machine-generated content. Users should be cautious and avoid blindly trusting the outputs of the LLMs without human verification. This problem could be potentially mitigated with the human-in-the-loop approach.

Finally, the use of automated tools for sustainability reporting raises concerns about transparency and accountability. It is crucial that the processes and algorithms used in our study are transparent and open to scrutiny. We commit to making our methods and data publicly available for review and validation by the broader community.

References

Marco Bronzini, Carlo Nicolini, Bruno Lepri, Andrea Passerini, and Jacopo Staiano. 2023. Glitter or gold? deriving structured insights from sustainability reports via large language models. *CoRR*, abs/2310.05628.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Salvatore Carta, Alessandro Giuliani, Leonardo Piano, Alessandro Sebastian Podda, Livio Pompianu, and Sandro Gabriele Tiddia. 2023. Iterative zero-shot LLM prompting for knowledge graph construction. *CoRR*, abs/2307.01128.

Harrison Chase. 2022. LangChain.

Nicola Guarino, Daniel Oberle, and Steffen Staab. 2009. What is an ontology? In Steffen Staab and Rudi Studer, editors, *Handbook on Ontologies*, International Handbooks on Information Systems, pages 1–17. Springer.

Aidan Hogan, Eva Blomqvist, Michael Cochez, Claudia d'Amato, Gerard de Melo, Claudio Gutierrez, Sabrina Kirrane, José Emilio Labra Gayo, Roberto Navigli, Sebastian Neumaier, Axel-Cyrille Ngonga Ngomo, Axel Polleres, Sabbir M. Rashid, Anisa Rula, Lukas Schmelzeisen, Juan Sequeda, Steffen Staab, and Antoine Zimmermann. 2021. *Knowledge Graphs*. Synthesis Lectures on Data, Semantics, and Knowledge. Morgan & Claypool Publishers.

Lars-Peter Meyer, Claus Stadler, Johannes Frey, Norman Radtke, Kurt Junghanns, Roy Meissner, Gordian Dziwis, Kirill Bulert, and Michael Martin. 2023. Llm-assisted knowledge graph engineering: Experiments with chatgpt. In *First Working Conference on Artificial Intelligence Development for a Resilient and Sustainable Tomorrow - AI Tomorrow 2023, Leipzig, Germany, 29-20 June, 2023*, Informatik Aktuell, pages 103–115. Springer.

Nandana Mihindukulasooriya, Sanju Tiwari, Carlos F. Enguix, and Kusum Lata. 2023. Text2kgbench: A benchmark for ontology-driven knowledge graph generation from text. In *The Semantic Web - ISWC 2023 - 22nd International Semantic Web Conference, Athens, Greece, November 6-10, 2023, Proceedings, Part II*, volume 14266 of *Lecture Notes in Computer Science*, pages 247–265. Springer.

Roberto Navigli, Paola Velardi, and Aldo Gangemi. 2003. Ontology learning and its application to automated terminology translation. *IEEE Intell. Syst.*, 18(1):22–31.

OpenAI. 2023. GPT-4 technical report. *CoRR*, abs/2303.08774.

Shirui Pan, Linhao Luo, Yufei Wang, Chen Chen, Jiapu Wang, and Xindong Wu. 2023. Unifying large language models and knowledge graphs: A roadmap. *CoRR*, abs/2306.08302.

Ridho Reinanda, Edgar Meij, and Maarten de Rijke. 2020. Knowledge graphs: An information retrieval perspective. *Found. Trends Inf. Retr.*, 14(4):289–444.

Milena Trajanoska, Riste Stojanov, and Dimitar Trajanov. 2023. Enhancing knowledge graph construction using large language models. *CoRR*, abs/2305.04676.

Anthony J Viera, Joanne M Garrett, et al. 2005. Understanding interobserver agreement: the kappa statistic. *Fam med*, 37(5):360–363.

Yuchen Zhou and Alexander Perzylo. 2023. Ontosustain: Towards an ontology for corporate sustainability reporting. In *Proceedings of the ISWC 2023 Posters, Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 22nd International Semantic Web Conference (ISWC 2023), Athens, Greece, November 6-10, 2023*, volume 3632 of *CEUR Workshop Proceedings*. CEUR-WS.org.

Yuqi Zhu, Xiaohan Wang, Jing Chen, Shuofei Qiao, Yixin Ou, Yunzhi Yao, Shumin Deng, Huajun Chen, and Ningyu Zhang. 2023. Llms for knowledge graph construction and reasoning: Recent capabilities and future opportunities. *CoRR*, abs/2305.13168.

A Appendix

TASK:Given the following ontology and sentences, please extract the triples from the sentence according to the relations in the ontology. In the output, only include given concepts, relations and the triples in the given output format. The resulting graph should only contain concept nodes from the ontology and strictly follow the relations between concepts. Do not add any attributes that do not appear in the schema shown below.

CONTEXT:

Ontology Concepts: Organization, Reporting, Impact, Risks, Opportunities, Strategy, Actions, Targets
Ontology Relations:(Organization, hasName, string),(Reporting, hasName, string),(Reporting, releaseDate, date),(Impact, hasDescription, string),(Risks, hasDescription, string),(Opportunities, hasDescription, string),(Strategy, hasDescriptiostn, string),(Actions, hasDescription, string),(Target, hasDescription, string)

Example Sentence: Siemens yearly report released for 2022 fiscal year. Due to our strategic focus on electrification, automation, and digitalization, we offer our customers highly efficient and longlived products that fulfill their function over a long period of time and are especially dependent on electricity for their operation. Our electric motors, which are efficient and long-lived, are an important factor in use phase emissions. Our targets are: to reduce emissions from business operations of Siemens without SHS by 55% by 2025 and 90% by 2030, and Net Zero operations by 2030 and supply chain by 2050. Use of biogas is another component of our decarbonization strategy, with which we have reduced our annual emissions by 19.6 thousand metric tons of CO2 compared to the use of conventional natural gas. We are working to reduce the emissions from our motor vehicle fleet, which comprises around 42,000 vehicles, and are striving to electrify it completely by 2030 as part of our EV100 commitment. We have increased the number of electric vehicles to around 1,360, and charging points to around 2,200. Landfill waste takes up space, generates greenhouse gas emissions, influences local biodiversity, and causes health problems for people and ecosystems. Where the environmentally sensitive use of energy is concerned, we deliberately go beyond the avoidance of emissions from power generation. Because even the generation of green electricity – for example, through the use of wind turbines or photovoltaics – has a negative impact on the environment since these systems have first to be manufactured, they change the local landscape when in operation, and they have to be disposed of at the end of their lifecycles. Society's rising expectations for corporate environmental responsibility have not only led to stricter legal regulations, but have also opened up new business opportunities, such as take back and recycling of products. Our intensified use of life cycle assessments and environmental product declarations enables us to identify environmentally compatible design alternatives that take circularity into account and can be integrated into product specifications.

Example output: (Organization, hasName, Siemens), (Reporting, hasName, Siemens report 2022), (Reporting, releaseDate, 2022), (Impact, hasDescription, Reduced greenhouse gas emissions\nImproved landfill waste management\nSecuring local biodiversity), (Risks, hasDescription, Product dependency on electricity may involve electricity price fluctuations, supply chain disruptions, or regulatory changes affecting energy usage\nFocus on decarbonization involves regulatory uncertainty, market volatility in renewable energy markets, and potential disruptions to supply chains), (Opportunities, hasDescription, Use biogas to reduce emissions\nTransition to electric vehicle\nCircular economy initiatives), (Strategy, hasDescription, Strategic focus on decarbonization, electrification, automation, and digitalization. We offer highly efficient and longlived products.Our electric motors are an important factor in use phase emissions), (Actions, hasDescription, Use of biogas reduced emissions by 19.6 thousand metric tons of CO2 compared to natural gas\nWorking on reducing emissions from motor vehicle fleet by electrifying it\nIncreased number of electric vehicles to 1360), (Targets, hasDescription, Emission reduction without SHS by 55% by 2025 and 90% by 2030\nNet Zero operations by 2030 and supply chain by 2050)

Test Sentence: Airbus SE is a European public company (Societas Europaea), with its seat in Amsterdam, the Netherlands, which is listed in ...
Test Output:

Figure 5: Prompt for knowledge extraction from Airbus report with relevant demonstration example from Siemens sustainability report 2022.

1	**Aspect**	**Topical match**	
2	**Organization**: Airbus		Annotator 1
3	**Report**: Airbus Registration Document 2021		
4	**Release date**: 2021		Please go through the extracted Topical standard aspects and evaluate them based on how well the extracted information matches the aspect descriptions. Evaluate each row and write Yes/No if you think the sentence is accurate, and it fits the description of the aspect.
5	**Strategy**: The company's strategy revolves around remaining competitive and financially resilient ...	yes	
6	**Impact**: - The company's operations and financial condition are negatively impacted...	no	
7	**Risks**: - The company faces substantial material risks due to climate change...	yes	
8	**Opportunities**: - Increasing regulatory pressure for energy transition, climate change ...	yes	
9	**Actions**: - The company implements risk management measures ...	no	
10	**Targets**: - It aims to introduce the world's first zero-emission commercial aircraft ...	yes	

Figure 6: An example form (Google Spreadsheets) featuring extracted Topical Standard aspects given to each annotator individually for annotation.

Strategy

- Strategy to stay competitive, invest in R&D, and manage diversified business portfolio in uncertain market and economic conditions
- Finances manufacturing activities and product development programmes through a combination of operating activities, customer advances, government advances and risk-sharing partnerships
- Bring a zero-emission aircraft to the market

Impact

- In the event of a systemic market disruption, the value and liquidity of the Company's financial instruments could decline resulting in significant impairment, negatively affecting the company's financial condition and operational results
- Potential increase in expenditure associated with environmental, human rights, health and safety regulations
- Impact on operating conditions of industrial activities due to climate change

Risks

- Financial markets remain unpredictable, which may cause the Company to increase its future outlays in connection with customer financing of commercial aircraft and helicopters
- Potential significant cash requirements related to COVID-19 crisis
- Financial instability in any part of the world can impact the company's ability to meet customer obligations

Opportunities

- Potential beneficial effects of the implementation of the Trade and Cooperation Agreement
- Trends in regulatory pressure indicate increasing demand for circular economy and resource efficiency, energy transition and climate change engagement, air and water quality improvement
- Energy transition, policy and Legal changes

Actions

- Ensure business practices conform to applicable laws, regulations and ethical business principles
- Management of logistical challenges due to travel limitations and restrictions
- Improved knowledge management and transfer schemes, seeking for development in sustainable technologies despite COVID-19's impact, use national research funding for ambitious programmes

Targets

- Aim to recover to pre-COVID levels between 2023 and 2025 in the commercial aircraft market
- Achievement of sustainability ambitions for future generations of aerospace through cross-industry and cross-government collaboration
- Efficient use of low carbon fuels for sustainable aviation. Zero-emission aircraft to the market

Figure 7: Example of extracted knowledge from Airbus annual report.