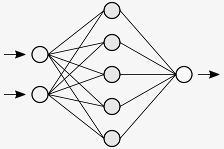
Open Universiteit

Leveraging Retrieval Augmented Generatıon for Literature Review

Ali D. Özbakır

19 December 2023 @ 11:00, MS Teams

Using the search query in TITLE - ABSTRACT - KEYWORD fields of journals, conference papers and book chapters



("machine learning" OR "artificial intelligence" OR "deep learning" OR "natural language processing")
AND
("greenwash*" OR "green claim" OR "green washing")
AND
(YEAR > 2017 AND LANGUAGE == ENG)

93

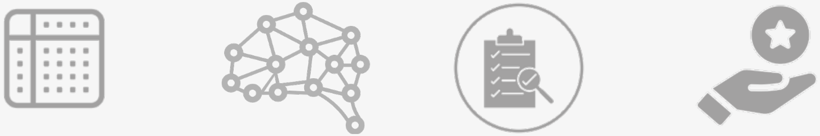
Round 1: Removed irrelevant, duplicate, short, inaccessible papers

30

Round 2: Read the IMRAD sections to ascertain greenwashing and AI

13

Round 3: Focus on extracting the following information:



It took 3 persons x 2 weeks, until the end of Round 2, without extensive domain knowledge

Environmental Claim Detection

Dominik Stammbach ETH Zurich dominsta@ethz.ch

Nicolas Webersinke FAU Erlangen-Nurnberg nicolas.webersinke@fau.de

Julia Anna Bingler Council on Economic Policies jb@cepweb.org

Mathias Kraus FAU Erlangen-Nurnberg mathias.kraus@fau.de

Markus Leippold University of Zurich markus.leippold@bf.uzh.ch

Own structure

Abstract

To transition to a green economy, environmental claims made by companies must be reliable, comparable, and verifiable. To analyze such claims at scale, automated methods are needed to detect them in the first place. However, there exist no datasets or models for this. Thus, this paper introduces the task of environmental claim detection. To accompany the task, we release an expert-annotated dataset and models trained on this dataset. We preview one potential application of such models: We detect environmental claims made in quarterly earning calls and find that the number of environmental claims has steadily increased since the Paris Agreement in 2015.

1 Introduction

In the face of climate change, we witness a transition towards a more sustainable and green economy. This change is driven by changes in regulation, public opinion, and investor attitudes. For example, global assets managed under a sustainability label are on track to exceed $53 trillion by 2025, more than a third of total assets under management. However, unfortunately, the boom has been accompanied by rampant greenwashing, with companies boasting about their environmental credentials.[1] Because of this surge in environmental claims and to protect consumers, initiatives on substantiating green claims are developed.[2] Due to an ever-growing amount of text, there is a need for automated methods to detect environmental claims. Detecting such claims at scale can assist policy-makers, regulators, journalists, activists, the research community, and an informed public in analyzing and scrutinizing environmental claims made by companies and facilitating the transition to a green economy.

Double columns

[1]See, e.g., The Economist, May 22nd, 2021.

[2]For example an EU initiative on green claims: https://ec.europa.eu/environment/eussd/smgp/initiative_on_green_claims.htm

Environmental claim: A total population of 6148 is getting the benefit of safe potable drinking water due to this initiative.

Environmental claim: Hydro has also started working on several initiatives to reduce direct CO2 emission in primary aluminium production.

Negative example: Generally, first of all, our Transmission department is very busy, both gas and electric transmission, I should say, meeting the needs of our on-network customers.

Negative example: Teams are thus focused on a shared objective in terms of growth and value creation.

Embedded text

Figure 1: Environmental Claims and Negative Examples from our dataset.

Thus, we introduce the task of environmental claim detection. Environmental claim detection is a sentence-level classification task with the goal of predicting whether a sentence contains an environmental claim or not. Often, environmental claims are made in a clear and concise matter on a sentence level, with the intention to convey to a consumer or stakeholder that a company or product is environmentally friendly.

To facilitate future research on environmental claim detection, we release an expert-annotated dataset containing real-world environmental claims and models which can be used by practitioners. For constructing the dataset, we were inspired by the European Commission (EC), which defines such claims as follows: *Environmental claims refer to the practice of suggesting or otherwise creating the impression (in the context of a commercial communication, marketing or advertising) that a product or a service is environmentally friendly (i.e., it has a positive impact on the environment) or is less damaging to the environment than competing goods or services.*[3] While such claims can be truthful and made in good faith, boasting about environmental credentials can also be monetized (de Freitas Netto

[3]From the Commission Staff Working Document, Guidance on the implementation/application of Directive 2005/29/EC on Unfair Commercial practices, Brussels, 3 December 2009 SEC(2009) 1666. See section 2.5 on misleading environmental claims.

 OOC

Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics
Volume 2: Short Papers, pages 1051–1066
July 9-14, 2023 ©2023 Association for Computational Linguistics

Source info here!

et al., 2020). For example, consumers are willing to spend more money on environmentally friendly products (Nielsen Media Research, 2015). The Commission states if environmental claims are too vague, unclear, or misleading, we are confronted with an instance of "greenwashing" (this definition is given in the same Commission Staff Working Document).

We situate environmental claim detection at the intersection of claim detection (e.g., Arslan et al., 2020) and pledge detection (Subramanian et al., 2019; Fornaciari et al., 2021). An environmental claim is typically made to increase the environmental reputation of a firm or a product. We show that models trained on the current claim and pledge detection datasets perform poorly at detecting environmental claims, hence the need for this new dataset. We make our dataset, code and models publicly available.[4] Lastly, we envision computer-assisted detection of greenwashing in future work, i.e., the automatic determination if an environmen-

Ethics Statement

Intended Use: This dataset will benefit journalists, activists, the research community, and an informed public analyzing environmental claims made by listed companies at scale. Also, we see this as a first step towards algorithmic greenwashing detection using NLP methods. It might also be useful to policy-makers and regulators in both the financial sector and the legal domain. Next, we hope companies are inspired by our work to produce more carefully drafted environmental claims. To conclude, we envision that the dataset and related models bring a large positive impact by encouraging truly environmentally friendly actions and less verbose boasting about environmental credentials.

Misuse Potential: Although we believe the intended use of this research is largely positive, there exists the potential for misuse. For example, it is possible that for-profit corporations will exploit AI models trained on this dataset while drafting

[10]https://www.sec.gov/

split	# examples	mean length	claims (%)
train	2117	24.4	0.25
dev	265	24.2	0.25
test	265	24.9	0.25
all	2647	24.5	0.25

Table 1: Dataset Statistics

Copy risk

Further, there exists work about claim verification of climate change related claims (Diggelmann et al., 2020), detecting media stance on global warming (Luo et al., 2020), collecting climate change opinions at scale from social platforms (Duong et al., 2022), and finally, the analysis of regulatory disclosures (Friederich et al., 2021; Kölbel et al., 2022).

truncated

In this broader context of applying NLP methods for climate change-related topics, We situate environmental claim detection at the intersection of claim spotting and pledge detection, covering the domain of text produced by companies with the

to participate in follow-up annotation work related to greenwashing detection.

different context

References

Ahmed Al-Rawi, Derrick O'Keefe, Oumar Kane, and Aimé-Jules Bizimana. 2021. Twitter's fake news discourses around climate change and global warming. *Front. Commun.*, 6.

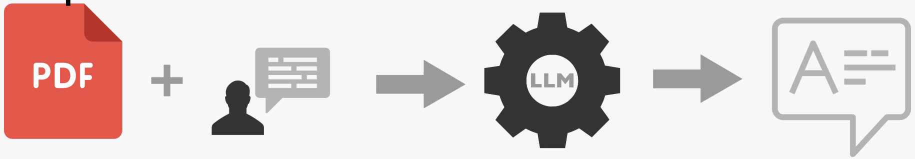
Fatma Arslan, Naeemul Hassan, Chengkai Li, and Mark Tremayne. 2020. A benchmark dataset of check-worthy factual claims. *Proceedings of the International AAAI Conference on Web and Social Media*, 14(1):821–829.

Pepa Atanasova, Alberto Barron-Cedeno, Tamer Elsayed, Reem Suwaileh, Wajdi Zaghouani, Spas Kyuchukov, Giovanni Da San Martino, and Preslav Nakov. 2018. Overview of the clef-2018 checkthat! lab on automatic identification and verification of political claims. task 1: Check-worthiness.

Alberto Barron-Cedeno, Tamer Elsayed, Preslav Nakov, Giovanni Da San Martino, Maram Hasanain, Reem Suwaileh, Fatima Haouari, Nikolay Babulkov, Bayan Hamdan, Alex Nikolov, Shaden Shaar, and Zien Sheikh Ali. 2020. Overview of checkthat! 2020:

Structure — Yet, copy/paste this document to a ChatBot and get ready for fun!

Prompt-based document retrieval



Redundancy - unnecessary repetition, diluting the response's relevance and overwhelming the reader.

Limited context - risks exceeding the model's context window, causing the loss of relevant information.

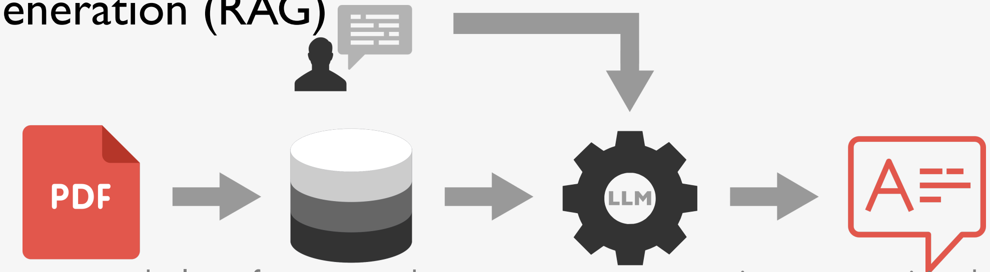
Processing overhead - potentially slowing down performance and consuming more resources..

Precision loss - the entire document might include irrelevant details that could potentially confuse the model..

Resource inefficiency - the entire document would require additional storage and computational resources.

Retrieval-augmented Generation (RAG)

LangChain
OpenAI
Chroma



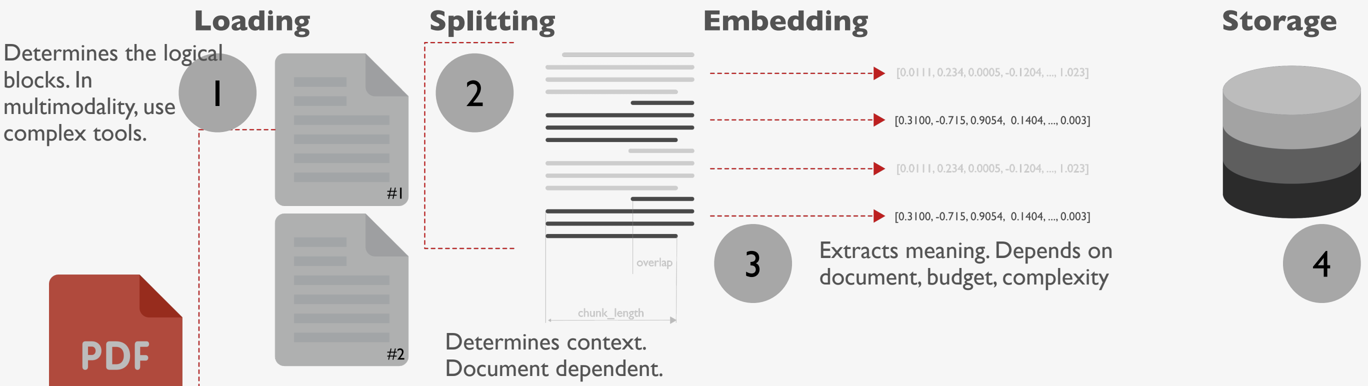
optimizes **relevance** by selectively incorporating pertinent information, avoiding unnecessary repetition.

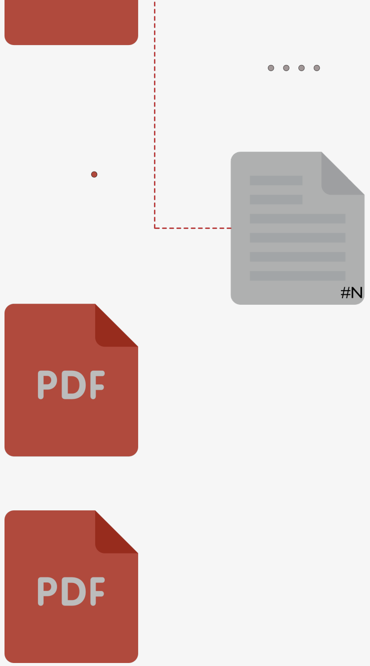
prevents the loss of pertinent information by dynamically fetching relevant details within the model's context window

enhances resource efficiency, minimizing processing overhead and resource consumption compared to copying entire documents.

improves precision by allowing the model to focus on relevant details, minimizing confusion caused by irrelevant information.

Resource inefficiency - the entire document would require additional storage and computational resources.



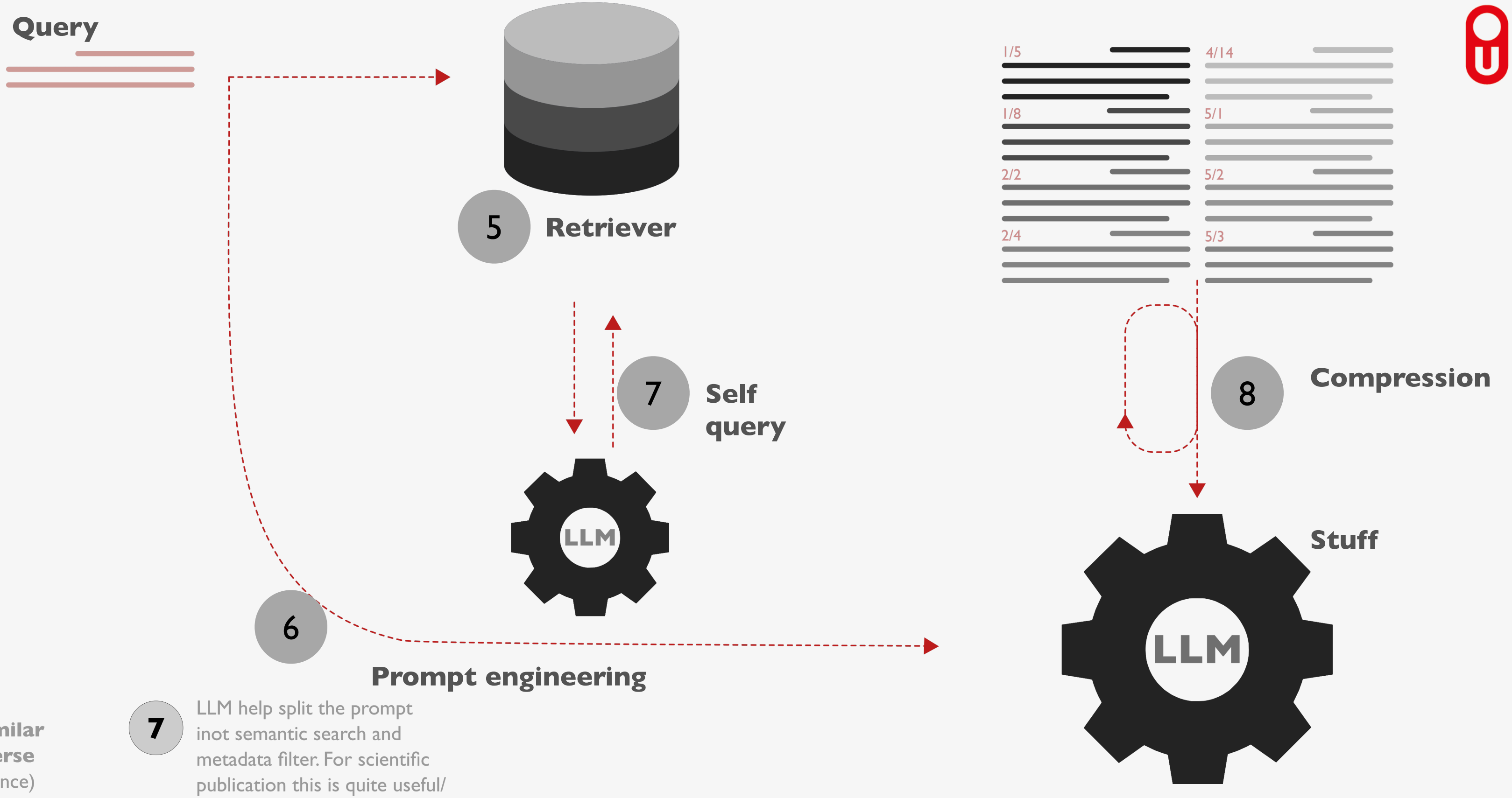


1. PyPDF, adds metadata={'source': dir, 'page': N}}. More complete metadata improves dense retrieval results, e.g. author, journal, title

2. First pre-process with .replace("- ", ""); RecursiveCharacterTextSplitter(chunk_length, chunk_overlap)

3. embeddings_model = OpenAIEmbeddings(
    model="text-embedding-ada-002")

4. Used ChromaDB to save vector embeddings. This is a light-weight solution with dense retrieval functions as well as metadata filtering. Also, there is a lot of examples to get inspiration :D

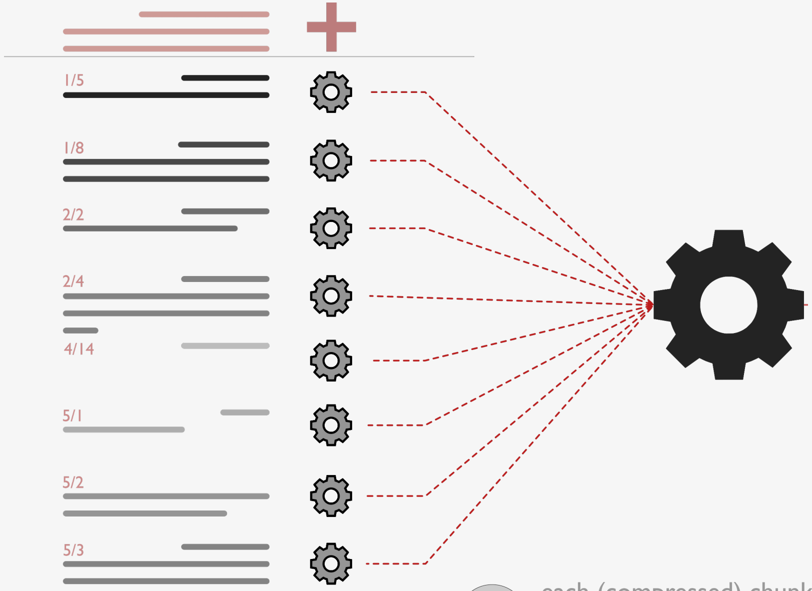


5 Find **n** similar chunks semantically most **similar** (Approx. NN) | **diverse** (max. Marginal Relevance)

6 Keep the query within context

7 LLM help split the prompt inot semantic search and metadata filter. For scientific publication this is quite useful/

8 LLM reduces the context documents to a few relevant sentences. Improtant to keep the num_tokens low.

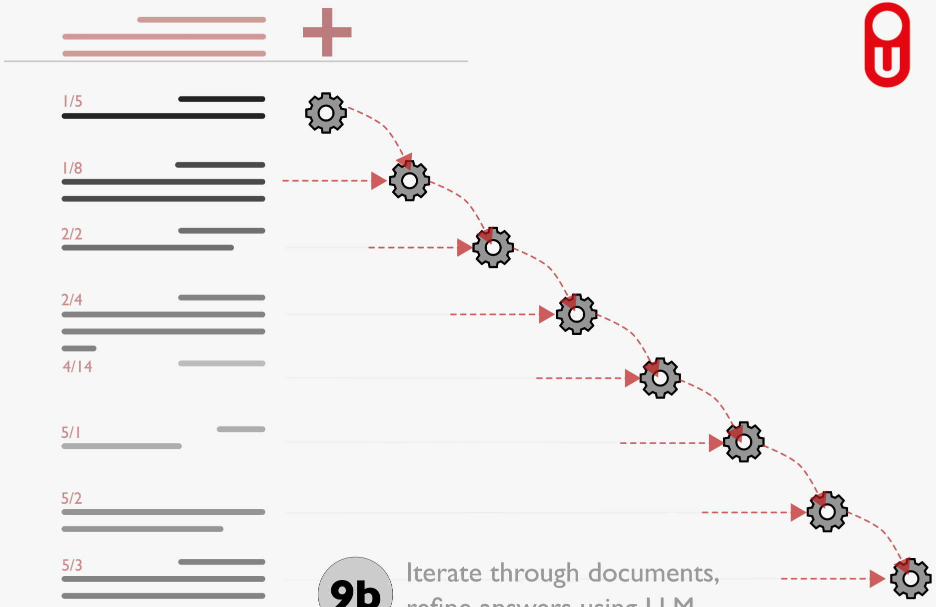


Map Reduce

9a

each (compressed) chunk is sent to the LLM to obtain an original answer, which subsequently are reduced to the final answer.

If the information we need is in multiple documents, we can check each one and use it to answer the question without worrying about space constraints in the context window.



Refine

9b

Iterate through documents, refine answers using LLM. More calls with more documents, cost increases. Outperforms map reduce for better results, continuous refinement.

Answer \ Retrieval	Incorrect	Correct
Correct	Chunking problem or embeddings model	
Incorrect	Search problem	LLM problem

Modified after Jarvis, C. OpenAI + Microsoft: Powering uour products with ChatGPT and your own data