

ChatClimate: Grounding Conversational AI in Climate Science

Veruska Muccione

University of Zurich <https://orcid.org/0000-0002-9773-3125>

Dominik Stammbach

ETH Zurich <https://orcid.org/0000-0003-1631-3020>

Jingwei Ni

ETH Zurich

Mathias Kraus

Friedrich-Alexander-Universität

Julia Bingler

University of Oxford

Simon Allen

University of Zurich <https://orcid.org/0000-0002-4809-649X>

Chiara Colesanti-Senni

University of Zurich

Tobias Wekhof

ETH Zurich <https://orcid.org/0000-0001-8564-2333>

Tobias Schimanski

University of Zurich

Glen Gostlow

University of Zurich

Nicolas Webersinke

Christian Huggel

University of Zurich <https://orcid.org/0000-0001-9299-2665>

Qian Wang

University of Zurich

Tingyu Yu

University of Zurich

Markus Leippold

markus.leippold@bf.uzh.ch

Saeid Vaghefi (✉ saeid.vaghefi@geo.uzh.ch)

University of Zurich <https://orcid.org/0000-0002-3061-6493>

Article

Keywords:

Posted Date: May 30th, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-2943500/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

ChatClimate: Grounding Conversational AI in Climate Science

Saeid Ashraf Vaghefi,^{1,2,3} Qian Wang,¹ Veruska Muccione,^{2,4} Dominik Stambach,⁵ Jingwei Ni,^{1,5} Mathias Kraus,⁶ Julia Bingler,^{7,8} Simon Allen,^{2,4} Chiara Colesanti-Senni,¹ Tobias Wekhof,^{1,5} Tobias Schimanski,¹ Glen Gostlow,¹ Tingyu Yu,¹ Nicolas Webersinke,⁶ Christian Huggel,² Markus Leippold^{1,9}

¹ Department of Banking and Finance, University of Zurich, Switzerland

² Department of Geography, University of Zurich, Switzerland

³ Eawag: Swiss Federal Institute of Aquatic Science, Switzerland

⁴ Institute for Environmental Sciences, University of Geneva, Switzerland

⁵ Center for Law and Economics, ETH Zurich, Switzerland

⁶ Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany

⁷ Smith School of Enterprise and the Environment, University of Oxford, Oxford, United Kingdom

⁸ Council on Economic Policies, Zurich, Switzerland

⁹ Swiss Finance Institute (SFI), Zurich, Switzerland

saeid.vaghefi@geo.uzh.ch, qian.wang@uzh.ch, veruska.muccione@geo.uzh.ch, dominik.stambach@gess.ethz.ch, njingwei@ethz.ch, mathias.kraus@fau.de, julia.bingler@smithschool.ox.ac.uk, simon.allen@geo.uzh.ch, chiara.colesantisenni@bf.uzh.ch, twekhof@ethz.ch, tobias.schimanski@bf.uzh.ch, glen.gostlow@bf.uzh.ch, tingyu.yu@bf.uzh.ch, nicolas.webersinke@fau.de, christian.huggel@geo.uzh.ch, markus.leippold@bf.uzh.ch

Abstract

Large Language Models (LLMs) have made significant progress in recent years, achieving remarkable results in question-answering tasks (QA). However, they still face two major challenges at inference time: hallucination and outdated information. These challenges take center stage in critical domains like climate change, where obtaining accurate and up-to-date information from reliable sources in a timely fashion is essential. To overcome these barriers, one potential solution is to provide LLMs with access to external, scientifically accurate, and robust sources (long-term memory). This access assists LLMs in continuously updating their knowledge and in preventing the propagation of inaccurate, incorrect, or outdated information. In this study, we enhanced GPT-4 by providing it access to the Sixth Assessment Report of the Intergovernmental (IPCC-AR6), the most comprehensive, up-to-date, and reliable source in this domain. We present our conversational AI prototype, available at www.chatclimate.ai,¹ and demonstrates its ability to answer challenging questions accurately in three different QA scenarios: asking from 1) GPT-4, 2) ChatClimate, and 3) hybrid ChatClimate. The answers and their sources were evaluated by a team of IPCC authors, who used their expert knowledge to score the accuracy of the answers from 1 (very-low) to 5 (very-high). The evaluation showed that the hybrid ChatClimate assistant provided more accurate answers, highlighting the effectiveness of our solution. This approach can be easily scaled to domain-specific chatbots, enabling the delivery of reliable and accurate information.

1 Introduction

Motivation. Large pre-trained Language Models (LLMs) have emerged as the de-facto standard in Natural Language Processing (NLP) in recent years. LLMs have revolutionized text processing across various tasks, bringing significant advancements in natural language understanding and generation (Devlin et al. 2019; Radford et al. 2019; Brown et al. 2020; Ouyang et al. 2022). Models such as LLaMA (Touvron et al. 2023), T0 (Sanh et al. 2021), PaLM (Chowdhery et al. 2022), GPT-3 (Brown et al. 2020), or instruction fine-tuned models, such as ChatGPT (OpenAI 2023b), GPT-4 (OpenAI 2023a) and HuggingGPT (Shen et al. 2023), have demonstrated exceptional capabilities in generating human-like text across various domains, including language translation, summarization, and question answering.

LLMs also excel at closed-book Question Answering (QA) tasks. Closed-book QA tasks require models to answer questions without any context (Li, Zhang, and Zhao 2022). LLMs like GPT-3/3.5 have achieved impressive results on multiple choice question answering (MCQA) tasks in the zero, one, and few-shot settings (Robinson, Rytting, and Wingate 2023). Recent works have used LLMs such as GPT-3 (Brown et al. 2020) as an implicit knowledge base which contains the necessary knowledge for answering questions (Shao et al. 2023).

However, LLMs suffer from two major issues: hallucination (Ji et al. 2023) and outdated information after training has concluded (Jang et al. 2022). These issues are particularly problematic in domains such as climate change, where it is critical to have accurate, reliable, and timely information.

¹under review.

We would like to express our gratitude to Francesco Leonetti from www.frigg.eco for his invaluable and voluntary support in setting up the server. The website has been available since mid-April.

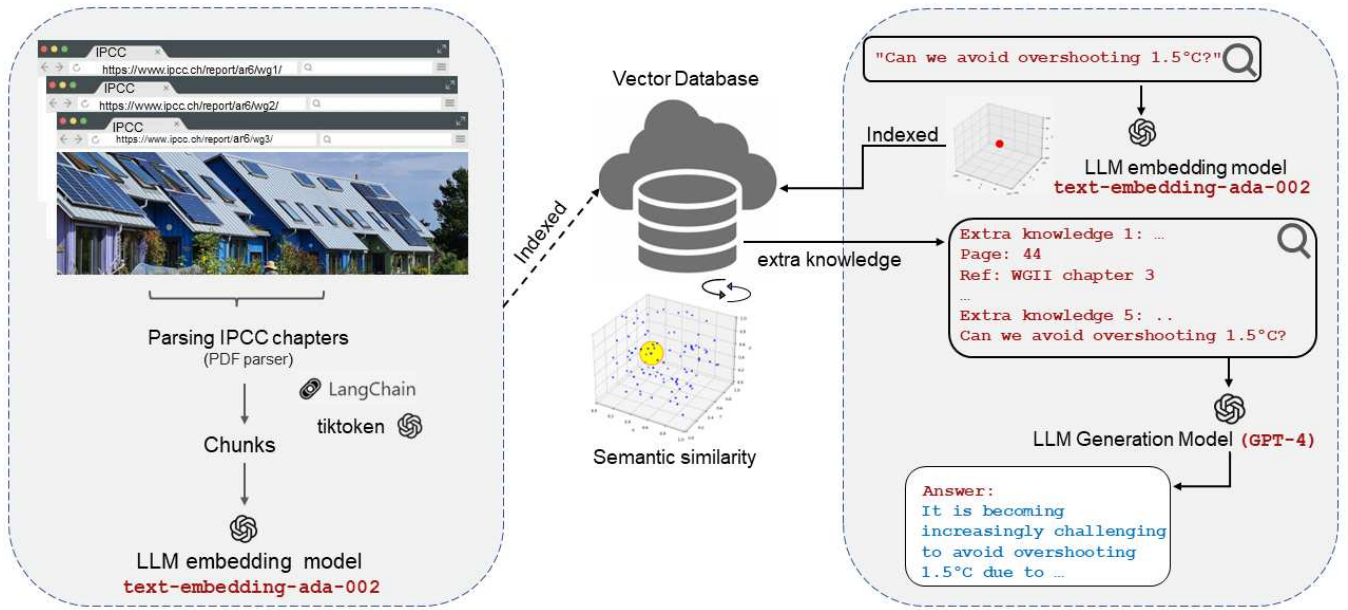


Figure 1: ChatClimate Data Pipeline: from creating external memory, receiving questions to accurate answers from IPCC AR6.

tion on changes in climate systems, current impacts, and projected risks of climate change and solution space. Hence, providing accurate and up-to-date responses with authoritative references and citations is paramount. Such responses, if accurate, can help understanding the scale and immediacy of the climate crisis, and facilitate the implementation of appropriate mitigation strategies.

Enhanced communication between government entities and the scientific community fosters more effective dialogue between national delegations and policymakers. A facilitated chat-based assisted feedback loop can be established by guaranteeing the accuracy of information sources and responses. This feedback loop promotes informed decision-making in relevant domains. For example, governments may ask a chatbot for feedback on specific statements in the report or request literature to support a claim. The importance of accurate and up-to-date information has been highlighted in previous studies as well (Bingler et al. 2022; Kumar, Singh, and Sethi 2021; Sethi, Singh, and Kumar 2020).

By overcoming the challenges of outdated information and hallucination, LLMs can be used to extract relevant information from large amounts of text and assist in decision-making. Training LLMs is computationally expensive and has other negative downsides (see e.g., Weidinger et al. 2021; Bender et al. 2021). To overcome the need for continuous training, one solution is to provide the LLMs with external sources of information (called Long-term memory). This memory continuously updates the knowledge of an LLM and reduces the propagation of incorrect or outdated information. Several studies have explored the use of external data sources, which makes the output of LLMs more factual (Borgeaud et al. 2022).

Contribution. In this paper, we introduce our prototype, ChatClimate (www.chatclimate.ai), a conversational AI de-

signed to improve the veracity and timeliness of LLMs in the domain of climate change by utilizing the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (hereafter IPCC AR6) (IPCC 2021, 2022a,b).² These reports offer the latest and most comprehensive evaluation of the climate system, climate change impacts, and solutions related to adaptation, mitigation, and climate-resilient development. We evaluate the LLMs’ performance in delivering accurate answers and references within the climate change domain by posing 13 challenging questions to our conversational AI (hereafter chatbot) across three scenarios: GPT-4, ChatClimate standalone, and a hybrid ChatClimate.

Findings. Our approach highlights the potential of incorporating external data sources to improve LLM performance in specific domains, such as climate change. By integrating up-to-date climate change information from the IPCC AR6 into LLMs, we can develop more accurate and reliable models for answering climate change-related questions. Our approach can potentially supply decision-makers and the public with trustworthy information on climate change, ultimately facilitating better-informed decision-making. This methodology can easily be scaled for chatbots in specialized domains to deliver more dependable and precise information.

The remainder of this paper is structured as follows. Section 2 gives a brief overview of recent developments in the field of LLMs, NLP in climate change, and the use of multiple sources in the context of LLMs. Section 3 describes the method with which we develop our prototype, and Section 4 showcases example questions that are asked to access in-

²As a disclaimer, please note that ChatClimate is not endorsed by the IPCC. ChatClimate will at times hallucinate and generate incorrect information. It may also occasionally produce harmful instructions or biased content.

formation from IPCC AR6. Section 5 discusses these results and the implications of this work. Section 6 concludes this work.

2 Background

Large Language Models

LLMs have transformed NLP and AI research over the last few years (Fan et al. 2023). They show surprising capabilities of generating creative text, solving basic math problems, answering reading comprehension questions, and much more. These models are based on the transformer architecture (Vaswani et al. 2017) and are trained on vast quantities of text data to identify patterns and connections within the data. Some notable examples of these models include GPT and BERT family models, which have been widely used for various NLP tasks (Devlin et al. 2019; Radford et al. 2019; Brown et al. 2020; Ouyang et al. 2022; OpenAI 2023a). The recent breakthroughs with models like T0 (Sanh et al. 2021), LLaMA (Touvron et al. 2023), PaLM (Chowdhery et al. 2022), GPT-3 (Brown et al. 2020), and GPT-4 (OpenAI 2023a) have further highlighted the potential of LLMs, with applications including chatbots (OpenAI 2023a,b) and virtual assistants (Jo 2023). However, LLMs suffer from hallucination (Ji et al. 2023), which refers to mistakes in the generated text that are simply made up or semantically incorrect. This can lead to vague or inaccurate responses to questions. Moreover, most of these models are trained on text prior to 2022, and thus have not been updated with new data or information since then (Ji et al. 2023).

NLP and Climate Change

NLP techniques have been widely used in the analysis of text related to climate change. Applications range from financial climate disclosure analyses (Bingler et al. 2022; Luccioni, Baylor, and Duchene 2020), detecting stance in media about global warming (Luo, Card, and Jurafsky 2020), detecting environmental claims (Stammach et al. 2022) and climate claims fact-checking (Diggelmann et al. 2020; Webersinke et al. 2022).

Question Answering and chatBots

Question-answering (QA) systems and chatbots have become increasingly popular. They can provide users with relevant and accurate information in a conversational setting. The importance, limitations, and future perspectives of conversational AI have been addressed in the literature from the open domain (OpenAI 2022; Adiwardana et al. 2020) to domain-specific chatbots (Lin et al. 2023). When presented with a question in human language, chatbots automatically provide a response. Although numerous information retrieval chatbots accomplish this task, transformer-based LLMs have become the de-facto standard in QA (Radford et al. 2019; Brown et al. 2020; OpenAI 2023a,b). In the context of climate change, QA systems and chatbots can help bridge the gap between complex scientific information and public understanding by providing concise and accessible answers to climate-related questions. Such systems

can also facilitate communication between experts, policy-makers, and stakeholders, enabling more informed decision-making and promoting climate change mitigation and adaptation strategies (Stammach et al. 2022; Callaghan et al. 2021). As the field of NLP and its application to climate change (Kölbel et al. 2020; Bingler et al. 2022) continues to advance, it is expected that QA systems and chatbots will play an increasingly important role in disseminating climate change information and fostering public engagement with climate science.

Long-term Memory and Agents for LLMs

One solution for enhancing the capabilities of LLMs in QA tasks is to fine-tune them on different datasets, which could be resource-wise expensive (Adiwardana et al. 2020). However, an alternative approach involves using agents that access the LLMs’ long-term memory, retrieve information, and insert it into a prompt to guide the LLMs more effectively (Kraus et al. 2023; Nair et al. 2023). These agents can decide which actions to perform, such as utilizing various tools, observing their outputs, or providing responses to user queries (Schick et al. 2023). This approach has been shown to improve the accuracy and efficiency of LLMs in a range of domains, including healthcare and finance (Schick et al. 2023). Domain-specific chatbots use a similar concept, where an agent has access to a carefully curated in-house database (long-term memory) to answer domain-specific questions (Gerhard-Young et al. 2022). These chatbots can provide customized responses based on the available information in their database, allowing for more accurate and relevant answers to user queries.

3 Method

ChatClimate Pipeline

In this study, we develop a long-term memory database by transforming the IPCC AR6 reports into a searchable format. The reports are converted from PDFs to JSON format, and each record in the database is divided into smaller chunks that LLM can easily process. The choice of the batch size for embeddings is a hyperparameter that requires tuning. We utilize OpenAI’s state-of-the-art text embedding model to vectorize each data chunk. Prior to injection into the database, we implement an efficient indexing mechanism to optimize retrieval times and facilitate effective information retrieval. Consequently, we can implement a semantic search that identifies the most relevant results based on the meaning and context of each query.

When a user poses a question, it is first embedded and then indexed using semantic similarity to find the five nearest vectors corresponding to the inquiry. The dot product of two vectors is utilized to analyze the similarity between vector embeddings, which is obtained by multiplying their respective components and summing the results.

After identifying the nearest vectors to the query vector, we retrieve the corresponding text from the database. The textual information is then used to refine and improve LLM prompts. Augmented queries are posed to the GPT-4 model

through instructed prompts, which enhance the user experience and increase the overall performance of our chatBot. Figure 1 shows the pipeline of ChatClimate.

Tools and External APIs

The first tool used in this study is a Python-based module that transforms IPCC AR6 reports from PDFs to JSON format (PDF parsr) and preprocesses the data, utilizing the powerful pandas library to access and manipulate data stored in dataframes.

The second tool is the LangChain Python package (<https://github.com/hwchase17/langchain>), which retrieves data from the JSON and chunks the extracted text into smaller sizes, ready for embedding. LangChain is a lightweight layer that transforms sequential LLM interactions into a natural conversation experience.

The third tool employed is OpenAI’s embedding model ”text-embedding-ada-002,” which vectorizes all chunks of the IPCC AR6 JSON files. Vector embeddings have proven to be a valuable tool for a variety of machine learning applications, as they can efficiently represent objects as dense vectors containing semantic information.

The fourth tool involves storing the generated vectors in a database, allowing for efficient storage and retrieval of the vector embeddings.

The fifth and final tool used is the GPT-4 ”chatcompletion” endpoint with instructed prompts, which provides answers to questions by leveraging the indexed vector embeddings.

Input Prompts and chatBots

The importance of prompt engineering for LLMs has been addressed in previous work (Schick and Schütze 2021; Reynolds and McDonell 2021; Kojima et al. 2023). We designed three prompts to compare the answers of our chatbots (i.e., ChatClimate, hybrid ChatClimate and GPT-4). The prompt used in our study consists of a series of instructions that guide the completion of a chat with GPT-4 on how to answer a provided question. The prompt is structured to allow the chatbot to access external resources while using its in-house knowledge.

Listing 1: Input prompt for hybrid ChatClimate.

```
"system", "content": You are a Q&A bot, an intelligent system that answers user questions based on the information provided by the user above the question and your in-house knowledge. There are five pieces of extra information above the user question. The user question is in the final line. When you use the user information, always indicate the Page and Reference in your answer. The Page and Reference are located below each piece of information. Additionally, let us know which part of your answer is from the IPCC information and which part is based on your in-house knowledge by writing either (IPCC AR6) or (In-house knowledge). If the information cannot be found in the information provided by
```

```
the user or your in-house knowledge, please say 'There is not enough information'.
"role": "user", "content": External knowledge + Question
```

Listing 2: Input prompt for ChatClimate.

```
"system", "content": You are a Q&A bot, an intelligent system that answers user questions ONLY based on the information provided by the user. There are five pieces of extra information above the user question. The user question is in the final line. When you use the user information, always indicate the Page and Reference in your answer. The Page and Reference are located below each piece of information. If the information cannot be found in the information provided by the user, please say 'There is not enough information'
"role": "user", "content": External knowledge + Question
```

Listing 3: Input prompt for GPT-4

```
"system", "content": "You are Q&A bot. A highly intelligent system that answers user questions"
"role": "user", "content": Question
```

Table 1: The 13 designed questions for running the 3 sets of experiments

Questions	Difficulty (1-5) 1:very-low, 2:low, 3:medium, 4:high, 5:very-high
Q1= "Is it still possible to limit warming to 1.5°C?"	3
Q2= "When will we reach 1.5°C?"	3
Q3= "What does overshoot mean?"	1
Q4= "Can we avoid overshooting 1.5°C?"	3
Q5= "Have emissions reductions fallen for some countries?"	3
Q6= "What are the issues with financing adaptation?"	4
Q7= "Where is the majority of climate finance going?"	4
Q8= "What are the options for scaling up adaptation and mitigation in developing countries?"	4
Q9= "Which regions will be disproportionately affected by climate change?"	5
Q10= "What is climate justice?"	2
Q11= "What is maladaptation?"	4
Q12= "Is there evidence of maladaptation?"	5
Q13= "Will glaciers in Scotland melt?"	5

Table 2: Comparison of generated answers to questions 1 and 2 from hybrid ChatClimate, ChatClimate, and GPT-4

hybrid ChatClimate	QA comparison	
	ChatClimate	GPT-4
Q1) Is it still possible to limit warming to 1.5°C?		
Yes, it is still possible to limit global warming to 1.5°C, but it would require immediate, deep, rapid, and sustained reductions in greenhouse gas emissions, including CO ₂ and methane (Page:31, IPCC AR6 WGIII Chapter03). Scenarios limiting warming to 1.5°C with no or limited overshoot show that net energy system CO ₂ emissions must fall by 35-51% by 2030 and 87-97% by 2050 (Page:43, IPCC AR6 WGIII Technical Summary). Net electricity sector CO ₂ emissions would need to reach zero globally between 2045 and 2055 (Page:3, IPCC AR6 WGIII Chapter06). However, the likelihood of limiting warming to 1.5°C with no or limited overshoot has decreased compared to the Special Report on Global Warming of 1.5°C due to the rise in global greenhouse gas emissions since its publication (Page:4, IPCC AR6 WGIII Chapter03). Only a small number of published pathways achieve this target without overshooting over the 21st century (Page:31, IPCC AR6 WGIII Technical Summary). In summary, while still possible, limiting warming to 1.5°C would require immediate and drastic action to reduce greenhouse gas emissions (In-house knowledge).	It is still possible to limit warming to 1.5°C with no or limited overshoot, but the likelihood has decreased since the Special Report on Global Warming of 1.5°C (SR1.5) due to the rise in global GHG emissions since it was published. This has led to higher near-term emissions (2030) and higher cumulative CO ₂ emissions until the time of net zero (medium confidence). Limiting warming to 1.5°C requires deep, rapid, and sustained reductions in greenhouse gas emissions, including methane, alongside rapid reductions of CO ₂ emissions to net zero IPCC AR6 WGIII Chapter03, Page:4, IPCC AR6 WGIII Technical Summary, Page:31.	Limiting global warming to 1.5°C is becoming increasingly challenging due to the continued rise in greenhouse gas emissions and the slow pace of implementing effective climate policies. However, it is not entirely impossible. Achieving this goal would require rapid and far-reaching transitions in energy, land, urban infrastructure, and industrial systems. This includes drastically reducing greenhouse gas emissions, increasing renewable energy sources, improving energy efficiency, and adopting sustainable land use practices. The Intergovernmental Panel on Climate Change (IPCC) has outlined various pathways to achieve this target, but it requires immediate and ambitious action from governments, businesses, and individuals worldwide.
accuracy: 5	accuracy: 5	accuracy: 3
Q2) When will we reach 1.5°C?		
We are likely to reach 1.5°C of global warming between 2030 and 2052 if the current rate of warming continues (high confidence) (IPCC AR6, Reference: IPCC AR6 WGI Chapter01, Page: 44). However, the exact timing may vary depending on the emissions reduction pathways and efforts to achieve net-zero CO ₂ emissions (In-house knowledge).	We are likely to reach 1.5°C of global warming between 2030 and 2052 if it continues to increase at the current rate (high confidence). Reference: IPCC AR6 WGI Chapter01, Page: 44.	As an AI, I cannot predict the exact year when global temperatures will reach 1.5°C above pre-industrial levels. However, according to the Intergovernmental Panel on Climate Change (IPCC), we could reach this threshold between 2030 and 2052 if global warming continues at its current rate. The actual timeline depends on various factors, including greenhouse gas emissions, climate policies, and global efforts to mitigate climate change.
accuracy: 5	accuracy: 4	accuracy: 2

hybrid ChatClimate: In this first scenario, the prompt starts with five pieces of external information retrieved from long-term memory, followed by a question that was asked by the user. The prompt instructs the chatbot to provide an answer based on the given information while using its own knowledge. Moreover, the chatbot is structured to prioritize IPCC AR6 for answers, referencing the names and pages of corresponding IPCC reports (Working Group I, II, III chapters, summary for policymakers, technical summary, and synthesis reports).

ChatClimate: In the second scenario, the prompt starts with five pieces of external information retrieved from long-term memory, followed by a question that was asked by the user. The prompt instructs the chatbot to provide answers only based on IPCC AR6.

GPT-4: In the last scenario, the prompt does not provide any extra information or instruction on how to provide answers and can be considered the baseline behaviour.

Overall, the prompt is designed to guide how to answer the questions given the availability of external and/or in-house knowledge. In the following, we demonstrate the three prompts used in this study: Listing 1, 2 and 3.

4 Experiments

chatBots and Questions

We conducted three sets of experiments by asking hybrid ChatClimate, ChatClimate, and GPT-4 chatbots 13 questions. Our team of IPCC AR6 authors then assessed the answers' accuracy. It's worth noting that our prototype's ability to provide sources for statements can facilitate the important pro-

Table 3: Cross-checking of the hybrid ChatClimate references for Question 9 with k-5, k-10, and k-15 nearest for extra knowledge retrieval

QA comparison		
Hybrid K-5	Hybrid K-10	Hybrid K-15
Q6) Where is the majority of climate finance going?		
<p>The majority of climate finance is going to a few large economies, with Brazil, India, China, and South Africa accounting for around one-quarter to more than a third depending on the year, a share similar to that represented by developed countries (IPCC AR6, WGIII Chapter 15, Page 18). Additionally, global climate finance is heavily focused on mitigation, with more than 90% on average between 2017-2020 being directed towards mitigation efforts (IPCC AR6, WGIII Chapter 15, Page 3).</p>	<p>Regions that will be disproportionately affected by climate change include East, Central, and West Africa, South Asia, Micronesia and Melanesia, and Central America (IPCC AR6, Page 81). These regions are highly vulnerable to climate hazards such as droughts, floods, and sea-level rise. In Asia, climate change hotspots with strong climate signals and high concentrations of vulnerable people are present in semi-arid, glacial-fed river basins and mega deltas (IPCC AR6, Page 10). South Asia and Africa are also identified as hotspots of multi-sectoral climate change risk, largely due to changes in water-related indicators affecting crop production (IPCC AR6, Page 81). Disproportionate impacts are exacerbated by factors such as poverty, hunger, conflict, and environmental degradation, which make countries and regions more vulnerable to climate change (IPCC AR6, Page 81). Vulnerable populations within these regions, including the poor, children, women, and marginalized Indigenous or ethnic minority people, often have limited resources and structures to support them during climate-related disasters (IPCC AR6, Page 27). Regions such as Africa and Central and South America are more likely to experience disproportionately higher impacts due to the increasing frequency of extreme temperatures, which can affect public health, agriculture, ecosystems, biodiversity, and infrastructure (IPCC AR6, Page 68). In addition, these regions are projected to face increased risks of flood, hunger, and water scarcity (IPCC AR6, Pages 75, 81, 110). In summary, regions that will be disproportionately affected by climate change include parts of Africa, South Asia, Central America, and specific climate-change hotspots in Asia. These regions face higher risks due to a combination of climate hazards, vulnerable populations, and limited resources for adaptation and support (In-house knowledge).</p>	<p>Regions that will be disproportionately affected by climate change include East, Central, and West Africa, South Asia, Micronesia and Melanesia, and Central America (IPCC AR6, WGII Chapter08, Page:81). These regions are highly vulnerable to climate hazards such as droughts, floods, and sea-level rise. In addition, Africa and Central and South America are more likely to experience disproportionately higher impacts due to their vulnerability, adaptation gaps, and development needs (IPCC AR6, WGII Chapter18, Page:68). South Asia and Africa are also identified as hotspots of multi-sectoral climate change risk, largely due to changes in water-related indicators affecting crop production (IPCC AR6, WGII Chapter16, Page:81). Climate change is expected to increase the magnitude and frequency of droughts and hurricanes in these regions, potentially forcing millions of people to leave their homes (IPCC AR6, WGII Chapter12, Page:80). Furthermore, vulnerable groups within these regions, such as the poor, children, women, and marginalized Indigenous or ethnic minority people, are differentially impacted by climate change (IPCC AR6, WGII Chapter08, Page:27). These vulnerable regions and countries often have limited resources and structures to support these groups when climate hazards occur, placing an additional burden on them (In-house knowledge). In summary, regions such as East, Central, and West Africa, South Asia, Micronesia and Melanesia, and Central America will be disproportionately affected by climate change due to their high vulnerability, exposure to climate hazards, and limited resources to support vulnerable populations (IPCC AR6, WGII Chapter08, Page:81).</p>
accuracy: 3	accuracy: 3	accuracy: 4

cess of trickle-back, which is often required by governments and other stakeholders in the context of IPCC reports.

Table 2 presents the returned answers from chatbots. The question "Is it still possible to limit warming to 1.5°C?" targets the mitigation, and the hybrid chatBot and ChatClimate explicitly return the greenhouse gas emission reduction amounts and time horizon while the GPT-4 answer is more general. All chatbots give a range of 2030 to 2052 as an answer to the question 'When will we reach 1.5°C?'. Hybrid chatbot and GPT-4 add to the answers that reaching 1.5°C, depends on the emission pathways

Evaluation of Answers (accuracy score)

Several studies focus on human-chatbot interaction effectiveness (Abdar, Tait, and Aleven 2020; Ramachandran, Es-

lami, and Sandvig 2020; Luger and Sellen 2016; Przegalin-ska et al. 2019). Evaluation involves factors such as relevance, clarity, tone, style, speed, consistency, personalization, error handling, and user satisfaction. This work, however, only examines the chatbot's performance on accuracy.

5 Discussion

Expert Cross-Check of the Answers

Overall, the responses provided by the hybrid ChatClimate were more accurate than those of ChatClimate standalone and GPT-4. For the sake of brevity, we have provided a detailed analysis of Q1 and Q2 in Table 2, and only highlight the key issues for Q3-Q13. For instance, in Q1, we asked the bots whether it is still possible to limit global warming

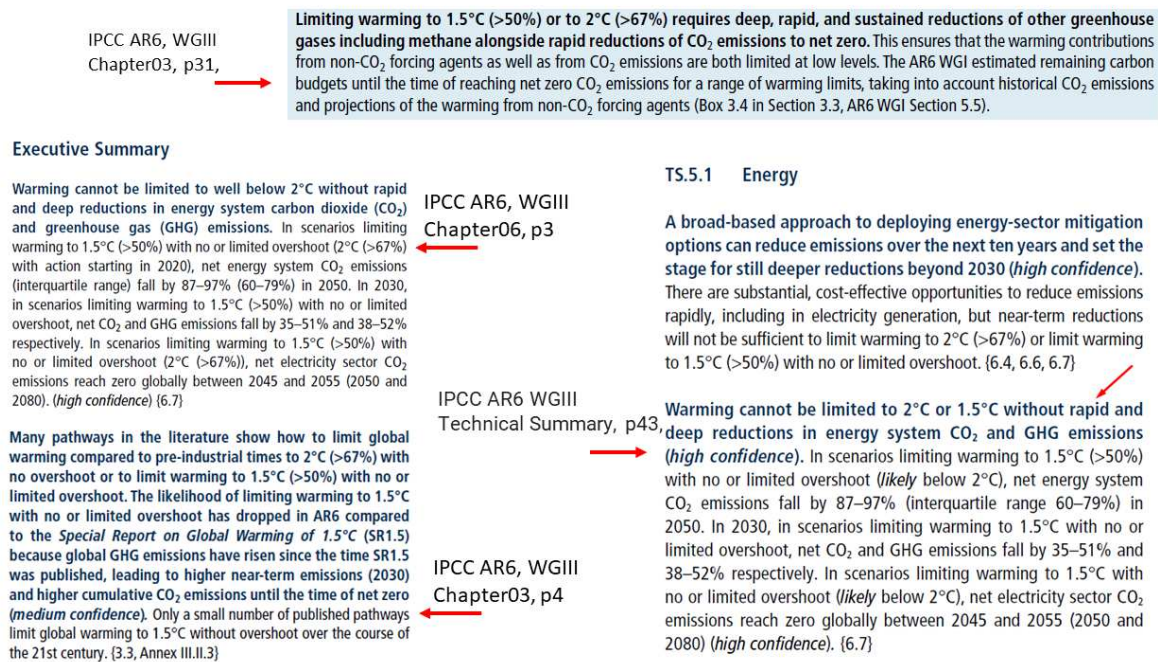


Figure 2: Cross-checking of the references for Question1

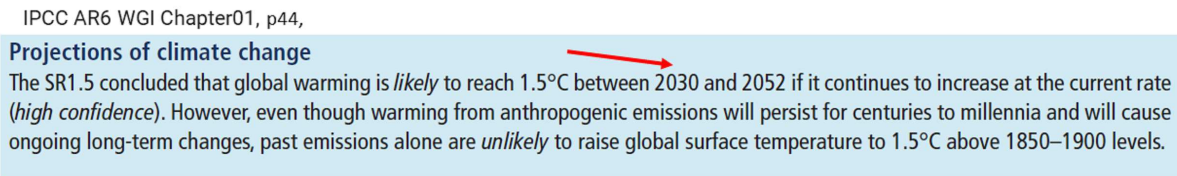


Figure 3: Cross-checking of the references for Question2

to 1.5°C. Both the hybrid ChatClimate and ChatClimate systems referred to the amount of CO₂ that needs to be reduced over different time horizons to stay below 1.5°C. However, the GPT-4 provided a more general response. To verify the accuracy of responses generated by the ChatClimate bots, we cross-checked the references provided by both systems. We found that the ChatClimate bots consistently provided sources for their statements, as shown in Figure 2 and 3, which is essential for verifying the veracity of the bot's answer. In Q2, we asked the bots about the time horizon when we reach 1.5°C. All three bots similarly referred to the time horizon of 2030 to 2052 based on the mitigation measures we take into account. The consistency of the answers shows that this time horizon has been mentioned in the training data of GPT-4 as well (e.g. IPCC AR6 WGI, which was released in August 2021 or Special Report of IPCC on Global Warming of 1.5 °C, 2018).

Prompts Engineering

By adjusting the retrieval hyperparameters, we can extract extra knowledge from long-term memory. We retrieved 10 and 15 nearest vectors to the user's query vector. With this,

GPT-4 can obtain more information and provide more accurate answers when accessing the outside world. The results have been demonstrated in Table 3, where K-15 scenario has provided more information about the regions that are most affected by climate change. This, in particular, shows the importance of prompt engineering,

Personalized-GPTs or GPT-n, Risk Management

Domain-specific chatbots and conversational AI tools provide easy access to accurate information. However, potential risks from external data sources, such as inaccuracies or biases, should be acknowledged. In this study, we developed and implemented domain-specific chatbots for the climate change domain. We compared three chatbot scenarios and found that the hybrid ChatClimate provided more accurate answers to 13 sample questions. We evaluated the answers internally, benefitting from the expert knowledge of co-authors. Since training LLMs is resource-intensive (OpenAI 2023a), integrating them with the outside world by providing long-term memory and prompt engineering could yield better results with fewer resources. However, creating long-term memory requires caution. We used the IPCC

AR6 as a comprehensive and reliable source to build external memory for LLMs, highlighting the importance of such databases for chatbot accuracy. Although there is an ongoing debate about pausing LLM training for six months until proper regulations are established, we believe that regulating LLM training, fine-tuning, and incorporating it into applications is necessary. Specifically, external database integration and prompt engineering should be considered in regulations for chatbots.

Limitations

chain of thoughts (COTs). In this study, we did not fully explore the potential of chain of thoughts (COTs) by testing different prompts. However, we expect that implementing COTs will improve the accuracy of our system’s outputs which we plan in our future works.

Evaluation of LLMs responses. We acknowledge that the evaluation of responses was not the primary focus of this work, and we relied solely on expert knowledge to assess the model’s performance. Additionally, further work is needed to provide a comprehensive description of the evaluation procedure, including aspects such as interannotator agreement and a more transparent explanation of query generation.

Future works

Fact-checking. Providing access for LLMs to trustworthy resources can enhance the model’s ability to perform fact-checking and provide well-grounded information to users. In ongoing research, we are exploring the potential of automated fact checking methods (e.g., Vlachos and Riedel 2014; Guo, Schlichtkrull, and Vlachos 2022). To this extent, we are building a authoritative and accurate knowledge base which can be used to fact-check domain-specific claims (Stammbach, Zhang, and Ash 2023) or LLM-produced responses. In this knowledge base, we will also leverage statements from the IPCC AR6 reports to validate or refute claims related to climate change and other environmental issues.

Further development. We continually improve ChatClimate and welcome community feedback on our website www.chatclimate.ai to enhance its question-answering capabilities. Our goal is to provide accurate and reliable climate change information, and we believe domain-specific chatbots like ChatClimate play a crucial role in achieving it.

6 Conclusion

In summary, our study demonstrates that the hybrid ChatClimate outperformed both GPT-4 and ChatClimate standalone in terms of the accuracy of answers when provided access to the outside world (IPCC AR6). The higher performance can be attributed to the integration of up-to-date and domain-specific data, which addresses the issues of hallucination and outdated information often encountered in LLMs. The results underline the importance of tailoring models to specific domains. The main findings of our work are summarized as follows:

1. The hallucination and outdated issues of LLMs could be refined by giving those models access to the knowledge beyond their training phase time and instructing LLMs on how to utilize that knowledge.
2. With prompt engineering and knowledge retrieval, LLMs can provide sources of the answers properly.
3. Hyperparameter tuning during knowledge retrieval and semantic search plays an important role in prompt engineering. We tested this by K-5, K-10, and K-15 nearest pieces of knowledge to the question in the semantic search between the question and the database.
4. Regulating LLM training, fine-tuning, and incorporating it into applications are necessary. Specifically, external database integration and prompt engineering should be considered in regulations for chatbots. We emphasize the importance of regulation for checking the outcomes of domain-specific chatbots. In such domains, users may not have enough knowledge to verify answers or cross-check references, making biased data or engineered prompts potentially harmful to end users.
5. Our findings not only emphasize the importance of leveraging climate domain information in QA tasks but also highlight the need for continued research and development in the field of AI-driven text processing.

References

- Abdar, M.; Tait, J.; and Aleven, V. 2020. The impact of chatbot characteristics on user satisfaction and conversational performance. *Journal of Educational Psychology*, 112(4): 667–683.
- Adiwardana, D.; Luong, M.-T.; So, D. R.; Hall, J.; Fiedel, N.; Thoppilan, R.; Yang, Z.; Kulshreshtha, A.; Nemade, G.; Lu, Y.; and Le, Q. V. 2020. Towards a Human-like Open-Domain Chatbot. arXiv:2001.09977.
- Bender, E. M.; Gebru, T.; McMillan-Major, A.; and Shmitchell, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, FAccT ’21, 610–623. New York, NY, USA: Association for Computing Machinery. ISBN 9781450383097.
- Bingler, J. A.; Kraus, M.; Leippold, M.; and Webersinke, N. 2022. Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures. *Finance Research Letters*, 102776.
- Borgeaud, S.; Mensch, A.; Hoffmann, J.; Cai, T.; Rutherford, E.; Millican, K.; van den Driessche, G.; Lespiau, J.-B.; Damoc, B.; Clark, A.; de Las Casas, D.; Guy, A.; Menick, J.; Ring, R.; Hennigan, T.; Huang, S.; Maggiore, L.; Jones, C.; Cassirer, A.; Brock, A.; Paganini, M.; Irving, G.; Vinyals, O.; Osindero, S.; Simonyan, K.; Rae, J. W.; Elsen, E.; and Sifre, L. 2022. Improving language models by retrieving from trillions of tokens. arXiv:2112.04426.
- Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell,

- A.; et al. 2020. Language models are few-shot learners. *Advances in neural information processing systems*, 33: 1877–1901.
- Callaghan, M.; Schleussner, C.-F.; Nath, S.; Lejeune, Q.; Knutson, T. R.; Reichstein, M.; Hansen, G.; Theokritoff, E.; Andrijevic, M.; Brecha, R. J.; et al. 2021. Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nature climate change*, 11(11): 966–972.
- Chowdhery, A.; Narang, S.; Devlin, J.; Bosma, M.; Mishra, G.; Roberts, A.; Barham, P.; Chung, H. W.; Sutton, C.; Gehrmann, S.; et al. 2022. Palm: Scaling language modeling with pathways. *arXiv preprint arXiv:2204.02311*.
- Devlin, J.; Chang, M.-W.; Lee, K.; and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805*.
- Diggelmann, T.; Boyd-Graber, J.; Bulian, J.; Ciaramita, M.; and Leippold, M. 2020. Climate-fever: A dataset for verification of real-world climate claims. *arXiv preprint arXiv:2012.00614*.
- Fan, L.; Li, L.; Ma, Z.; Lee, S.; Yu, H.; and Hemphill, L. 2023. A Bibliometric Review of Large Language Models Research from 2017 to 2023. *arXiv:2304.02020*.
- Gerhard-Young, G.; Anantha, R.; Chappidi, S.; and Hoffmeister, B. 2022. Low-Resource Adaptation of Open-Domain Generative Chatbots. *arXiv:2108.06329*.
- Guo, Z.; Schlichtkrull, M.; and Vlachos, A. 2022. A Survey on Automated Fact-Checking. *Transactions of the Association for Computational Linguistics*, 10: 178–206.
- IPCC. 2021. Climate Change 2021: The Physical Science Basis. *Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- IPCC. 2022a. Climate Change 2022: Impacts, Adaptation, and Vulnerability. *Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- IPCC. 2022b. Climate Change 2022: Mitigation of Climate Change. *Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*.
- Jang, J.; Ye, S.; Yang, S.; Shin, J.; Han, J.; Kim, G.; Choi, S. J.; and Seo, M. 2022. Towards Continual Knowledge Learning of Language Models. In *ICLR*.
- Ji, Z.; Lee, N.; Frieske, R.; Yu, T.; Su, D.; Xu, Y.; Ishii, E.; Bang, Y. J.; Madotto, A.; and Fung, P. 2023. Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys*, 55(12): 1–38.
- Jo, A. 2023. The Promise and Peril of Generative AI. *Nature*, 614.
- Kojima, T.; Gu, S. S.; Reid, M.; Matsuo, Y.; and Iwasawa, Y. 2023. Large Language Models are Zero-Shot Reasoners. *arXiv:2205.11916*.
- Kölbel, J. F.; Leippold, M.; Rillaerts, J.; and Wang, Q. 2020. Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure. *Available at SSRN 3616324*.
- Kraus, M.; Bingler, J. A.; Leippold, M.; Schimanski, T.; Senni, C. C.; Stambach, D.; Vaghefi, S. A.; and Webersinke, N. 2023. Enhancing Large Language Models with Climate Resources. *arXiv:2304.00116*.
- Kumar, A.; Singh, S.; and Sethi, N. 2021. Climate change and cities: challenges ahead. *Frontiers in Sustainable Cities*, 3: 645613.
- Li, J.; Zhang, Z.; and Zhao, H. 2022. Self-Prompting Large Language Models for Open-Domain QA. *arXiv:2212.08635*.
- Lin, B.; Bouneffouf, D.; Cecchi, G.; and Varshney, K. R. 2023. Towards Healthy AI: Large Language Models Need Therapists Too. *arXiv:2304.00416*.
- Luccioni, A.; Baylor, E.; and Duchene, N. 2020. Analyzing Sustainability Reports Using Natural Language Processing. *arXiv:2011.08073*.
- Luger, E.; and Sellen, A. 2016. Towards a framework for evaluation and design of conversational agents. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2885–2891. ACM.
- Luo, Y.; Card, D.; and Jurafsky, D. 2020. Detecting Stance in Media On Global Warming. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, 3296–3315. Online: Association for Computational Linguistics.
- Nair, V.; Schumacher, E.; Tso, G.; and Kannan, A. 2023. DERA: Enhancing Large Language Model Completions with Dialog-Enabled Resolving Agents. *arXiv:2303.17071*.
- OpenAI. 2022. ChatGPT: A Large-scale Generative Language Model for Conversational AI.
- OpenAI. 2023a. GPT-4 Technical Report. Technical report, OpenAI.
- OpenAI. 2023b. InstructGPT: AI for Generating Instructions. <https://openai.com/research/instructgpt/>.
- Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35: 27730–27744.
- Przegalinska, A.; Ciechanowski, L.; Stroz, A.; Gloor, P.; and Mazurek, G. 2019. In bot we trust: A new methodology of chatbot performance measures. *Business Horizons*, 62(6): 785–797. Digital Transformation and Disruption.
- Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; and Sutskever, I. 2019. Language Models are Unsupervised Multitask Learners. *OpenAI Blog*.
- Ramachandran, D.; Eslami, M.; and Sandvig, C. 2020. A Framework for Understanding and Evaluating Automated Systems. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 154–164.
- Reynolds, L.; and McDonell, K. 2021. Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI EA '21. New York, NY, USA: Association for Computing Machinery. ISBN 9781450380959.

- Robinson, J.; Rytting, C. M.; and Wingate, D. 2023. Leveraging Large Language Models for Multiple Choice Question Answering. *arXiv:2210.12353*.
- Sanh, V.; Webson, A.; Raffel, C.; Bach, S. H.; Sutawika, L.; Alyafeai, Z.; Chaffin, A.; Stiegler, A.; Scao, T. L.; Raja, A.; Dey, M.; Bari, M. S.; Xu, C.; Thakker, U.; Sharma, S. S.; Szczechla, E.; Kim, T.; Chhablani, G.; Nayak, N.; Datta, D.; Chang, J.; Jiang, M. T.-J.; Wang, H.; Manica, M.; Shen, S.; Yong, Z. X.; Pandey, H.; Bawden, R.; Wang, T.; Neeraj, T.; Rozen, J.; Sharma, A.; Santilli, A.; Fevry, T.; Fries, J. A.; Teehan, R.; Biderman, S.; Gao, L.; Bers, T.; Wolf, T.; and Rush, A. M. 2021. Multitask Prompted Training Enables Zero-Shot Task Generalization. *arXiv:2110.08207*.
- Schick, T.; Dwivedi-Yu, J.; Dessì, R.; Raileanu, R.; Lomeli, M.; Zettlemoyer, L.; Cancedda, N.; and Scialom, T. 2023. Toolformer: Language models can teach themselves to use tools. *arXiv preprint arXiv:2302.04761*.
- Schick, T.; and Schütze, H. 2021. It’s Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners. *arXiv:2009.07118*.
- Sethi, N.; Singh, S.; and Kumar, A. 2020. The importance of accurate and up-to-date information in the context of climate change. *Journal of Cleaner Production*, 277: 123304.
- Shao, Z.; Yu, Z.; Wang, M.; and Yu, J. 2023. Prompting Large Language Models with Answer Heuristics for Knowledge-based Visual Question Answering. *arXiv:2303.01903*.
- Shen, Y.; Song, K.; Tan, X.; Li, D.; Lu, W.; and Zhuang, Y. 2023. HuggingGPT: Solving AI Tasks with ChatGPT and its Friends in HuggingFace. *arXiv:2303.17580*.
- Stammach, D.; Webersinke, N.; Bingler, J. A.; Kraus, M.; and Leippold, M. 2022. A Dataset for Detecting Real-World Environmental Claims. *arXiv preprint arXiv:2209.00507*.
- Stammach, D.; Zhang, B.; and Ash, E. 2023. The Choice of Textual Knowledge Base in Automated Claim Checking. *J. Data and Information Quality*, 15(1).
- Touvron, H.; Lavril, T.; Izacard, G.; Martinet, X.; Lachaux, M.-A.; Lacroix, T.; Rozière, B.; Goyal, N.; Hambro, E.; Azhar, F.; et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, L. u.; and Polosukhin, I. 2017. Attention is All you Need. In Guyon, I.; Luxburg, U. V.; Bengio, S.; Wallach, H.; Fergus, R.; Vishwanathan, S.; and Garnett, R., eds., *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc.
- Vlachos, A.; and Riedel, S. 2014. Fact Checking: Task definition and dataset construction. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, 18–22. Baltimore, MD, USA: Association for Computational Linguistics.
- Webersinke, N.; Kraus, M.; Bingler, J. A.; and Leippold, M. 2022. ClimateBert: A Pretrained Language Model for Climate-Related Text. *arXiv:2110.12010*.
- Weidinger, L.; Mellor, J.; Rauh, M.; Griffin, C.; Uesato, J.; Huang, P.-S.; Cheng, M.; Glaese, M.; Balle, B.; Kasirzadeh, A.; Kenton, Z.; Brown, S.; Hawkins, W.; Stepleton, T.; Biles, C.; Birhane, A.; Haas, J.; Rimell, L.; Hendricks, L. A.; Isaac, W.; Legassick, S.; Irving, G.; and Gabriel, I. 2021. Ethical and social risks of harm from Language Models. *arXiv:2112.04359*.

Appendix

Expert cross-check of the QA scenarios Q3-13

Here we continue the demonstration of QA and our analyses for questions 3 to 13. It is important to note that hallucination issues are more effectively addressed by the ChatClimate standalone version, which draws answers exclusively from IPCC AR6. This is evident in instances such as Q12, addressing maladaptation, and Q13, concerning melting glaciers. In Q12, GPT-4 offers answers related to maladaptation across various sectors. Meanwhile, in Q13, both GPT-4 and the hybrid produce hallucinations since there is no explicit statement regarding glaciers in Scotland. In these cases, we anticipate a definitive answer akin to the one provided by the ChatClimate standalone. We believe automated fact-checking methods (e.g., Vlachos and Riedel 2014; Guo, Schlichtkrull, and Vlachos 2022) can further help address hallucinations in a post-processing step where we first perform an automated fact-check of the chatbot response before returning it to the user. To this extent, we are in the process of developing an authoritative and accurate knowledge base that can be used to fact-check domain-specific climate change related LLM responses (Stammach, Zhang, and Ash 2023). The knowledge base will be a collection of timely peer-reviewed scientific articles and the IPCC AR6 and similar reports. We hope that this approach can help minimize the amount of hallucinations produced by a chat-based assistant.

Table SM1: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question3

hybrid ChatClimate	ChatClimate	GPT-4
Q3) What does overshoot mean?		
Overshoot refers to the exceedance of a specified global warming level, followed by a decline to or below that level during a specified period of time (e.g., before 2100). In the context of climate change, overshoot pathways are scenarios where a certain concentration, forcing, or global warming level is first exceeded and then returns to or below that level before the end of a specified period of time. The overshoot duration can vary, but in most cases, it occurs over a period of at least one decade and up to several decades (IPCC AR6 WGI AnnexVII, Page:28; IPCC AR6 WGIII Annex I, Page:20).	Overshoot refers to the exceedance of a specified global warming level, followed by a decline to or below that level during a specified period of time (e.g., before 2100). Sometimes the magnitude and likelihood of the overshoot are also characterized. The overshoot duration can vary from one pathway to the next, but in most overshoot pathways in the literature and referred to as overshoot pathways in the AR6, the overshoot occurs over a period of at least one decade and up to several decades. See also Pathways. Reference: IPCC AR6 WGII Annex-II, Page: 29.	Overshoot refers to a situation where a value or process exceeds its intended target or goal, often temporarily, before stabilizing or returning to the desired level. In various contexts, it can refer to: <ol style="list-style-type: none"> 1. In ecology, overshoot occurs when a population exceeds the carrying capacity of its environment, leading to a decline in resources and potential population collapse. 2. In economics, overshoot can refer to a situation where economic growth or inflation exceeds its sustainable level, leading to a subsequent correction or downturn. 3. In engineering and control systems, overshoot refers to the transient response of a system that exceeds its final steady-state value before settling down to that value. 4. In physics, overshoot can refer to a situation where an object moves beyond its intended stopping point, often due to momentum or external forces, before returning to the desired position.
accuracy: 5	accuracy: 4	accuracy: 2

Table SM2: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question4

hybrid ChatClimate	ChatClimate	GPT-4
Q4) Can we avoid overshooting 1.5°C?		
<p>Avoiding overshooting 1.5°C is challenging, but not impossible. The likelihood of limiting warming to 1.5°C with no or limited overshoot has dropped in AR6 compared to the Special Report on Global Warming of 1.5°C (SR1.5) because global GHG emissions have risen since the time SR1.5 was published, leading to higher near-term emissions (2030) and higher cumulative CO2 emissions until the time of net zero (medium confidence) (IPCC AR6 WGIII Chapter03, Page:4). Only a small number of published pathways limit global warming to 1.5°C without overshoot over the course of the 21st century (IPCC AR6 WGIII Chapter03, Page:4).</p> <p>Mitigation pathways with 2030 emissions levels in the NDCs announced prior to COP26 consistently overshoot 1.5°C by 0.15°C-0.3°C (IPCC AR6 WGIII Chapter03, Page:60). To avoid overshooting 1.5°C, it would require rapid and significant reductions in greenhouse gas emissions, as well as the implementation of effective climate policies and technologies (In-house knowledge).</p>	<p>Based on the information provided, avoiding overshooting 1.5°C is challenging. The likelihood of limiting warming to 1.5°C with no or limited overshoot has dropped in AR6 compared to the Special Report on Global Warming of 1.5°C (SR1.5) because global GHG emissions have risen since the time SR1.5 was published, leading to higher near-term emissions (2030) and higher cumulative CO2 emissions until the time of net zero (medium confidence). Only a small number of published pathways limit global warming to 1.5°C without overshoot over the course of the 21st century (IPCC AR6 WGIII Chapter03, Page:4).</p>	<p>As an AI, I cannot predict the future, but I can provide information on the current situation. Avoiding overshooting 1.5°C of global warming above pre-industrial levels is a challenging task, as we have already experienced around 1.1°C of warming. To avoid overshooting 1.5°C, rapid and far-reaching actions in energy, land use, urban planning, transportation, and industrial systems are required. This would involve transitioning to renewable energy, increasing energy efficiency, reducing deforestation, and adopting sustainable agriculture practices, among other measures. The Intergovernmental Panel on Climate Change (IPCC) has outlined various pathways to achieve this goal, but it requires immediate and ambitious efforts from governments, businesses, and individuals worldwide.</p>
accuracy: 5	accuracy: 4	accuracy: 3

Table SM3: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question5

hybrid ChatClimate	ChatClimate	GPT-4
Q5) Have emissions reductions fallen for some countries?		
<p>Yes, emissions reductions have fallen for some countries. A growing number of countries have reduced CO₂ and GHG emissions for longer than 10 years (high confidence). Data up to 2018 indicates that about 24 countries have reduced territorial CO₂ and GHG emissions (excluding LULUCF CO₂), as well as consumption-based CO₂ emissions, for at least 10 years (Page 22, IPCC AR6 WGIII Chapter02). At least 18 countries have sustained production-based GHG and consumption-based CO₂ emission reductions for longer than 10 years (Page 11, IPCC AR6 WGIII SummaryForPolicymakers; Page 13, IPCC AR6 WGIII SPM). These reductions were linked to energy supply decarbonisation, energy efficiency gains, and energy demand reduction, which resulted from both policies and changes in economic structure (Page 11, IPCC AR6 WGIII SummaryForPolicymakers; Page 13, IPCC AR6 WGII SPM). However, the total reduction in annual GHG emissions of these countries is small (about 3.2 GtCO₂-eq per yr) compared to global emissions growth observed over the last decade (Page 3, IPCC AR6 WGIII Chapter02).</p> <p>accuracy: 5</p>	<p>Yes, a growing number of countries have achieved GHG emission reductions for longer than 10 years. Data up to 2018 indicates that about 24 countries have reduced territorial CO₂ and GHG emissions (excluding LULUCF CO₂), as well as consumption-based CO₂ emissions, for at least 10 years. At least 18 countries have sustained production-based GHG and consumption-based CO₂ emission reductions for longer than 10 years. These reductions were linked to energy supply decarbonization, energy efficiency gains, and energy demand reduction, which resulted from both policies and changes in economic structure. However, these reductions have only partly offset global emissions growth (Reference: IPCC AR6 WGIII Chapter02, Page:22; IPCC AR6 WGIII SummaryForPolicymakers, Page:11; IPCC AR6 WGIII SPM, Page:13; IPCC AR6 WGIII Chapter02, Page:3).</p> <p>accuracy: 4</p>	<p>Yes, emissions reductions have fallen for some countries over the years. Various factors contribute to this, such as the implementation of climate policies, increased use of renewable energy sources, improvements in energy efficiency, and economic shifts. The extent of emissions reductions varies across countries, depending on their specific circumstances and efforts to combat climate change. However, it is important to note that global greenhouse gas emissions continue to rise, and more collective action is needed to achieve significant reductions on a global scale.</p> <p>accuracy: 3</p>

Table SM4: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question6

hybrid ChatClimate	ChatClimate	GPT-4
Q6) What are the issues with financing adaptation?		
<p>There are several issues with financing adaptation, including:</p> <ol style="list-style-type: none"> 1. Implicit economic world views and technical support availability: The choice of funding and financing mechanisms is often based on implicit economic world views or the technical support available to sub-national governments, such as preparing municipal bonds or contracting for public-private partnerships (IPCC AR6, WGII Chapter 6, Page 70). 2. Justice implications: Adaptation finance has profound justice implications, as it can intensify financial and environmental risks borne primarily by the poor, the working class, or people discriminated against because of race, sexual orientation, or ability (IPCC AR6, WGII Chapter 6, Page 70). 3. Dwindling resources and trade-offs: Discussions of adaptation finance often occur in the context of dwindling resources and trade-offs, leading to triage decisions about other investments that societies can tolerate suspending (IPCC AR6, WGII Chapter 8, Page 78). 4. Governance challenges: In many poor, vulnerable countries, complex governance challenges, such as budget austerity or corruption, hamper the provision of support for adaptation (IPCC AR6, Chapter 8, Page 78). 5. Insufficient financing: Insufficient financing is a key driver of adaptation gaps, with annual finance flows targeting adaptation being significantly less than the lowest adaptation cost estimates for near-term climate change (IPCC AR6, WGII Technical Summary, Page 50). 6. Limited private-sector financing: Globally, private-sector financing of adaptation has been limited, especially in developing countries (IPCC AR6 WGII, Technical Summary, Page 50). 7. Weak incentives for climate-resilient financial strategies: The use of climate-resilient financial strategies and instruments is limited by weak incentives, high upfront costs, high transaction and intermediation costs, and relatively long pay-off times (IPCC AR6, WGII Chapter 5, Page 113). 8. Imbalanced allocation of public and private finance: Given limited experience and weak information for adaptation finance, sub-optimal outcomes may include imbalanced allocation of public and private finance, as well as inequitable division of risks and returns (IPCC AR6, WGII Chapter 5, Page 113). 	<p>There are several issues with financing adaptation:</p> <ol style="list-style-type: none"> 1. Implicit economic world views and technical support availability influencing the choice of funding and financing mechanisms (Page 70, IPCC AR6 WGII Chapter06). 2. Adaptation finance having profound justice implications, as it can intensify financial and environmental risks primarily borne by the poor, working class, or people discriminated against due to race, sexual orientation, or ability (Page 70, IPCC AR6 WGII Chapter06). 3. Dwindling resources and trade-offs, leading to triage decisions about other investments that societies can tolerate suspending (Page 78, IPCC AR6 WGII Chapter08). 4. Complex governance challenges, such as budget austerity or corruption, hampering the provision of support in many poor, vulnerable countries (Page 78, IPCC AR6 WGII Chapter08). 5. In the absence of adaptation funding for the poor, the costs of adaptation are borne by the poor at community, kin-group, and household scales, which can ultimately increase the likelihood of remaining trapped in poverty (Page 78, IPCC AR6 WGII Chapter08). 6. Insufficient financing being a key driver of adaptation gaps, with finance not targeting more vulnerable countries and communities (Page 50, IPCC AR6 WGII TechnicalSummary) 7. Limited use of climate-resilient financial strategies and instruments due to weak incentives, high upfront costs, high transaction and intermediation costs, and relatively long pay-off times (Page 113, IPCC AR6 WGII Chapter05). 8. Imbalanced allocation of public and private finance, as well as inequitable division of risks and returns (Page 113, IPCC AR6 WGII Chapter05). 	<p>Financing adaptation to climate change involves several challenges and issues, including:</p> <ol style="list-style-type: none"> 1. Insufficient funding: The amount of funding available for adaptation projects is often inadequate to meet the growing needs of countries and communities vulnerable to climate change impacts. 2. Uncertainty in climate change impacts: The unpredictable nature of climate change makes it difficult to accurately estimate the costs of adaptation measures, which can lead to underfunding or misallocation of resources. 3. Prioritization of adaptation projects: With limited funding, it is challenging to prioritize adaptation projects based on their urgency, effectiveness, and potential benefits, leading to potential inefficiencies in resource allocation. 4. Lack of access to finance: Developing countries, particularly the least developed and most vulnerable, often face difficulties in accessing international climate finance due to complex procedures, lack of capacity, and limited resources. 5. Inadequate integration of adaptation into development planning: Adaptation measures are often not well-integrated into national and local development plans, leading to a lack of coherence and coordination in addressing climate change impacts. 6. Limited private sector involvement: The private sector has a crucial role to play in financing adaptation, but barriers such as perceived risks, lack of incentives, and limited awareness of opportunities often hinder their engagement. 7. Short-term focus: Funding for adaptation is often focused on short-term projects, while long-term planning and investments are needed to build resilience to climate change. 8. Monitoring and evaluation challenges: Assessing the effectiveness of adaptation projects is difficult due to the lack of standardized metrics, methodologies, and data, which can hinder learning and improvement in adaptation planning and financing. 9. Capacity constraints: Many developing countries lack the technical, institutional, and financial capacity to plan, implement, and monitor adaptation projects effectively. 10. Equity concerns: Financing for adaptation should be distributed fairly, taking into account the needs and vulnerabilities of different countries and communities. However, this can be challenging due to competing priorities, limited resources, and political considerations.
accuracy: 4	accuracy: 4	accuracy: 3

Table SM5: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question7

QA comparison		
hybrid ChatClimate	ChatClimate	GPT-4
Q6) Where is the majority of climate finance going?		
<p>The majority of climate finance is going to a few large economies, with Brazil, India, China, and South Africa accounting for around one-quarter to more than a third depending on the year, a share similar to that represented by developed countries (IPCC AR6, WGIII Chapter 15, Page 18). Additionally, global climate finance is heavily focused on mitigation, with more than 90% on average between 2017-2020 being directed towards mitigation efforts (IPCC AR6, WGIII Chapter 15, Page 3).</p>	<p>The majority of climate finance is going to a few large economies, with Brazil, India, China, and South Africa accounting for around one-quarter to more than a third depending on the year, a share similar to that represented by developed countries. Least-developed countries (LDCs) continue to represent less than 5% year-on-year (IPCC AR6 WGIII Chapter15, Page:18). Additionally, global climate finance is heavily focused on mitigation, with more than 90% on average between 2017-2020 being directed towards mitigation efforts (IPCC AR6 WGIII Chapter15, Page:3).</p>	<p>The majority of climate finance is directed towards mitigation and adaptation efforts to address climate change. Mitigation efforts focus on reducing greenhouse gas emissions, promoting renewable energy, and enhancing energy efficiency. Adaptation efforts aim to help communities and countries adapt to the impacts of climate change, such as investing in resilient infrastructure, improving water management, and supporting sustainable agriculture. Climate finance is distributed through various channels, including public and private sources, international climate funds, and development banks. The specific allocation of funds may vary over time and depend on the priorities and needs of different regions and sectors.</p>
accuracy: 4	accuracy: 4	accuracy: 3

Table SM6: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question8

hybrid ChatClimate	ChatClimate	GPT-4
Q8) What are the options for scaling up adaptation and mitigation in developing countries?		
Options include:	Options for scaling up adaptation and mitigation in developing countries include:	There are several options for scaling up adaptation and mitigation in developing countries. These options can be broadly categorized into the following areas:
<ol style="list-style-type: none"> 1. Increased levels of public finance and publicly mobilized private finance flows from developed to developing countries in the context of the USD100 billion-a-year goal (IPCC AR6 WGIII SPM, Page 51). 2. Increase the use of public guarantees to reduce risks and leverage private flows (IPCC AR6 WGIII SPM, Page 51). 3. Scaled-up public grants for mitigation and adaptation funding for vulnerable regions, especially in Sub-Saharan Africa, which would be cost-effective and have high social returns in terms of access to basic energy (IPCC AR6 WGIII SPM, Page 51). 4. Exploiting available mechanisms for expanding the adaptive capacity of human and natural systems (IPCC AR6 WGII Chapter08, Page 57). 5. Implementing widely applicable adaptation options that could be scaled up to reduce vulnerability or exposure for the majority of the world's population and the ecosystems they depend on, such as nature restoration, changing diets and reducing food waste, infrastructure retrofitting, building codes, disaster early warning, and cooperative governance (IPCC AR6 WGII Chapter17, Page 4). 	<ol style="list-style-type: none"> 1. Accelerated financial support from developed countries and other sources to enhance mitigation action and address inequities in access to finance (IPCC AR6 WGIII SPM, Page:51). 2. Scaled-up public grants for mitigation and adaptation funding for vulnerable regions, especially in Sub-Saharan Africa, which would be cost-effective and have high social returns in terms of access to basic energy (IPCC AR6 WGIII SPM, Page:51). 3. Increased levels of public finance and publicly mobilized private finance flows from developed to developing countries in the context of the USD 100 billion-a-year goal (IPCC AR6 WGIII SPM, Page:51). <p>Additionally, some widely applicable adaptation options that could be scaled up to reduce vulnerability or exposure include nature restoration, changing diets and reducing food waste, infrastructure retrofitting, building codes, disaster early warning, and cooperative governance (IPCC AR6 WGII Chapter17, Page:4). However, the specific portfolio of adaptation options may vary across locations and contexts.</p>	<ol style="list-style-type: none"> 1. Capacity building and technical assistance: Developing countries often lack the necessary knowledge, skills, and resources to implement effective adaptation and mitigation measures. Providing technical assistance, training, and capacity building can help these countries better understand and address climate change challenges. 2. Financial support: Scaling up adaptation and mitigation efforts requires significant financial resources. Developed countries can provide financial support through grants, loans, and other mechanisms to help developing countries implement climate change projects and programs. 3. Technology transfer: Access to advanced technologies can significantly enhance the ability of developing countries to adapt to climate change and reduce greenhouse gas emissions. Technology transfer can be facilitated through partnerships, licensing agreements, and other collaborative arrangements. 4. Policy and regulatory frameworks: Developing countries can benefit from the development and implementation of robust policy and regulatory frameworks that promote climate-resilient development and low-carbon growth. This may include policies related to land use, energy, transportation, and agriculture, among others. 5. Public-private partnerships: Engaging the private sector in adaptation and mitigation efforts can help mobilize resources and expertise. Public-private partnerships can be an effective way to scale up climate action in developing countries. 6. Knowledge sharing and networking: Sharing best practices, lessons learned, and successful models can help developing countries learn from each other and scale up their efforts. Regional and international networks can facilitate knowledge sharing and collaboration. 7. Mainstreaming climate change into development planning: Integrating climate change considerations into national and sectoral development plans can help ensure that adaptation and mitigation measures are prioritized and adequately funded. 8. Monitoring and evaluation: Regular monitoring and evaluation of adaptation and mitigation efforts can help identify successes, challenges, and areas for improvement. This information can be used to inform future efforts and scale up effective strategies.
These options are based on both IPCC information and in-house knowledge.		
accuracy: 4	accuracy: 4	accuracy: 3

Table SM7: Cross-checking of the references for Question9

QA comparison		
hybrid GPT-4 ChatClimate	ChatClimate	GPT-4
Q9) Which regions will be disproportionately affected by climate change?		
<p>The regions that will be disproportionately affected by climate change include East, Central, and West Africa, South Asia, Micronesia and Melanesia, and Central America (IPCC AR6, Page 81). These regions are highly vulnerable to climate hazards such as droughts, floods, and sea level rise. In addition, regions like Africa and Central and South America will experience a disproportionate exposure to temperature extremes, exacerbating their vulnerability, adaptation gaps, and development needs (IPCC AR6, Page 68). Climate change hotspots in Asia, such as semi-arid areas, glacial-fed river basins, and mega deltas, also face strong climate signals and high concentrations of vulnerable people (IPCC AR6, Page 10).</p>	<p>The regions that will be disproportionately affected by climate change include East, Central, and West Africa, South Asia, Micronesia and Melanesia, and Central America. These regions are highly vulnerable to climate change impacts and have a higher number of people exposed to climate hazards such as droughts, floods, or sea-level rise (Reference: IPCC AR6 WGII Chapter08, Page: 81). Additionally, regions like Africa and Central and South America experience disproportionate exposure due to their vulnerability, adaptation gaps, and development needs (Reference: IPCC AR6 WGII Chapter18, Page: 68).</p>	<p>Regions that are expected to be disproportionately affected by climate change include:</p> <ol style="list-style-type: none"> 1. Low-lying coastal areas: These regions are at a higher risk of flooding and storm surges due to rising sea levels. Examples include small island nations, the Netherlands, and coastal cities like Miami and New York. 2. Polar regions: The Arctic and Antarctic are experiencing rapid warming, leading to melting ice caps and permafrost, which can have global implications for sea levels and weather patterns. 3. Sub-Saharan Africa: This region is particularly vulnerable to climate change due to its reliance on rain-fed agriculture, high levels of poverty, and limited adaptive capacity. Climate change is expected to exacerbate water scarcity, food insecurity, and the spread of diseases like malaria. 4. South and Southeast Asia: Countries like Bangladesh, India, and the Philippines are vulnerable to climate change impacts such as sea-level rise, extreme weather events, and increased temperatures, which can lead to reduced agricultural productivity, water scarcity, and health issues. 5. Small island developing states: These nations, such as those in the Caribbean and Pacific, are highly vulnerable to sea-level rise, storm surges, and the loss of coral reefs, which can impact tourism, fisheries, and coastal infrastructure. 6. Central and South America: Climate change can exacerbate existing environmental challenges, such as deforestation and water scarcity, and lead to more frequent and severe weather events, impacting agriculture and infrastructure. 7. Australia: The continent is already experiencing more frequent and severe heatwaves, droughts, and bushfires, which can have significant impacts on ecosystems, agriculture, and human health. <p>It is important to note that climate change impacts are interconnected and can have cascading effects across regions and sectors. Vulnerability to climate change is influenced by factors such as socioeconomic status, governance, and access to resources, which can vary significantly within and between regions.</p>
accuracy: 3	accuracy: 3	accuracy: 4

Table SM8: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question10

hybrid ChatClimate	QA comparison	
	ChatClimate	GPT-4
Q10) What is climate justice?		
Climate justice is a term used to address the ethical and political aspects of global warming by examining inequalities and violation of human collective rights in relation to climate change impacts (IPCC AR6 WGII Chapter 10, Page 75). It focuses on the disproportionate impacts of climate change on the most vulnerable populations and calls for community sovereignty and functioning (IPCC AR6 WGIII Chapter 14, Page 58). Climate justice involves protecting vulnerable populations from climate change impacts, mitigating the effects of transformations, and easing the transition for those whose livelihoods rely on fossil fuel-based sectors (IPCC AR6 WGIII Chapter 17, Page 42). It also considers international justice and fairness between states, emphasizing the principle of common but differentiated responsibilities (CBDR) and respective capabilities, which reflects the idea that all countries must address climate change, but the form of climate action depends on their situation (IPCC AR6 WGII Chapter 01, Page 42).	Climate justice is a term used to address the disproportionate impacts of climate change on the most vulnerable populations. It involves examining inequalities and violation of human collective rights in relation to climate-change impacts by employing the concepts of environmental justice and social justice. Climate justice calls for community sovereignty and functioning, and focuses on issues such as unequal global power relations, wealth, interests within communities, and historical responsibility for climate change. In international climate negotiations, the principle of 'common but differentiated responsibilities' is often used to balance these issues, reflecting the idea that all countries must address climate change, but the form of climate action depends on their situation and capacities (IPCC AR6 WGIII Chapter14, Page:58; IPCC AR6 WGII Chapter01, Page:42).	Climate justice refers to the concept that addresses the ethical and political aspects of climate change, emphasizing the need for equitable distribution of responsibilities, resources, and opportunities related to climate change mitigation and adaptation. It acknowledges that the impacts of climate change disproportionately affect vulnerable and marginalized communities, who often contribute the least to greenhouse gas emissions. Climate justice seeks to promote social, economic, and environmental equity, ensuring that all people have the right to a healthy environment and a sustainable future. This includes advocating for policies and actions that reduce emissions, protect ecosystems, and support the resilience and adaptation of affected communities.
accuracy: 4	accuracy: 4	accuracy: 4

Table SM9: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question11

QA comparison		
hybrid ChatClimate	ChatClimate	GPT-4
Q11) What is maladaptation?		
Maladaptation refers to current or potential negative consequences of adaptation-related responses that lead to an increase in the climate vulnerability of a system, sector, or group by exacerbating or shifting vulnerability or exposure now or in the future, and eroding sustainable development (IPCC AR6 WGII Chapter 17, Page 62). It involves actions that may lead to increased risk of adverse climate-related outcomes, including increased greenhouse gas emissions, increased or shifted vulnerability to climate change, more inequitable outcomes, or diminished welfare, now or in the future. Most often, maladaptation is an unintended consequence (IPCC AR6 WGII Summary for Policymakers, Page 7).	Maladaptation refers to actions that may lead to increased risk of adverse climate-related outcomes, including increased greenhouse gas emissions, increased or shifted vulnerability to climate change, more inequitable outcomes, or diminished welfare, now or in the future. Most often, maladaptation is an unintended consequence. It refers to current or potential negative consequences of adaptation-related responses that lead to an increase in the climate vulnerability of a system, sector, or group by exacerbating or shifting vulnerability or exposure now or in the future, and eroding sustainable development (IPCC AR6 WGII SummaryForPolicymakers, Page:7; IPCC AR6 WGII Chapter17, Page:62).	Maladaptation refers to a trait, behavior, or response that is not well-suited or poorly adapted to a particular situation or environment. It can negatively impact an individual's ability to cope with or adjust to specific circumstances, often leading to decreased functionality, inefficiency, or harm. Maladaptation can occur in various contexts, such as biological, psychological, or social, and can result from genetic factors, learned behaviors, or environmental influences.
accuracy: 5	accuracy: 5	accuracy: 1

Table SM10: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question12

hybrid ChatClimate	QA comparison	
	ChatClimate	GPT-4
Q12) Is there evidence of maladaptation?		
Yes, there is increasing evidence of maladaptation globally (IPCC AR6 WGII Chapter16, Page:26). Maladaptation has been particularly reported in the context of agricultural, forestry and fisheries practices, migration in the Global South, and some infrastructure-based interventions. Urban heat adaptations have been linked to maladaptation that increase health risks and/or energy (IPCC AR6 WGII Chapter16, Page:26). In some sectors and systems, inappropriate responses to climate change create long-term lock-in of vulnerability, exposure, and risks that are difficult and costly to change (IPCC AR6 WGII TechnicalSummary, Page:51). These maladaptations can exacerbate existing inequalities for Indigenous Peoples and vulnerable groups, impeding the achievement of Sustainable Development Goals (SDGs) and increasing adaptation needs (IPCC AR6 WGII TechnicalSummary, Page:51).	Yes, there is increasing evidence of maladaptation globally (high confidence). Maladaptation has been particularly reported in the context of agricultural, forestry and fisheries practices, migration in the Global South, and some infrastructure-based interventions. Urban heat adaptations have been linked to maladaptation that increase health risks and/or energy (IPCC AR6 WGII Chapter16, Page:26).	<p>Yes, there is evidence of maladaptation in various species, including humans. Maladaptation refers to a trait or behavior that is more harmful than helpful to an organism's survival and reproduction. It can occur due to various reasons, such as rapid environmental changes, genetic mutations, or human intervention. Some examples of maladaptation include:</p> <ol style="list-style-type: none"> 1. The peacock's tail: While the large and colorful tail of the male peacock is attractive to females and helps in mating, it also makes the bird more visible to predators and hinders its ability to escape. 2. The dodo bird: The dodo bird was a flightless bird that lived on the island of Mauritius. Its lack of fear towards humans and inability to fly made it an easy target for hunters, ultimately leading to its extinction. 3. Antibiotic resistance: The overuse of antibiotics in humans and animals has led to the evolution of antibiotic-resistant bacteria. This maladaptation has made it increasingly difficult to treat bacterial infections, posing a significant threat to public health. 4. Human lifestyle diseases: Modern human lifestyles have led to an increase in sedentary behavior and unhealthy diets, contributing to the rise of obesity, diabetes, and cardiovascular diseases. These maladaptive behaviors can negatively impact an individual's health and well-being. 5. Climate change: Human activities, such as the burning of fossil fuels and deforestation, have led to rapid climate change. This has resulted in maladaptation for many species, as they struggle to adapt to the changing environment, leading to population declines and even extinction for some species.
accuracy: 5	accuracy: 4	accuracy: 1

Table SM11: Comparison of generated answers from hybrid ChatClimate, ChatClimate, and GPT-4 to question13

hybrid ChatClimate	QA comparison	
	ChatClimate	GPT-4
Q13) Will glaciers in Scotland melt?		
Yes, it is very likely that glaciers in Scotland will melt. Glaciers are projected to continue to lose mass under all emissions scenarios (very high confidence) and this includes glaciers in various regions, such as the Arctic and Antarctic (high confidence) (IPCC AR6 WGI Chapter08, Page:68). Although the specific mention of Scotland is not provided in the provided references, the overall trend of glacier melting applies to glaciers worldwide, including those in Scotland (In-house knowledge).	There is not enough information provided to specifically answer the question about glaciers in Scotland.	Glaciers in Scotland, like those in other parts of the world, are susceptible to melting due to climate change and rising global temperatures. While Scotland currently has only a few small glaciers and permanent snowfields, they could potentially melt or shrink significantly in the future if global warming continues at its current pace. The rate and extent of melting will depend on various factors, including future greenhouse gas emissions, regional climate patterns, and local topography.
accuracy: 3	accuracy: 5	accuracy: 1