

---

# AI and Climate Information Needs in Africa

---

Afra Amini<sup>†‡</sup> Hannah Brown<sup>♣</sup> Christian Buck<sup>†</sup> Jannis Bulian<sup>†</sup>  
Michelle Chen Huebscher<sup>†</sup> Massimiliano Ciaramita<sup>†\*</sup> Sonali Das<sup>♣</sup> Ben Gaiarin<sup>††</sup>  
Christopher Gordon<sup>◇</sup> Rangan Gupta<sup>♣</sup> Kenny Kutu<sup>♣</sup> Doreen Lartey<sup>◇</sup> Markus Leippold<sup>†•</sup>  
Leslie Leuenberger<sup>††</sup> Mark Atta Mensah<sup>◇</sup>

<sup>‡</sup> ETH Zürich, <sup>††</sup> Google, <sup>†</sup> Google DeepMind,  
<sup>◇</sup> University of Ghana, <sup>♣</sup> University of Pretoria, <sup>•</sup> University of Zurich

## Abstract

What are Africa’s climate information needs? How do they relate to global concerns? How can AI help? This paper begins exploring these topics, with a specific focus on Ghana and South Africa. Our study involves climate-related questions from Google Trends and identifies macro clusters related to the broad themes of understanding climate change, its causes, and potential solutions. A finer-grained analysis reveal country-specific topical profiles, including pronounced concerns, in the two African countries, about the disruptive consequences on land and economy. We also begin to probe the ability of large language models to answer some of these questions. Preliminary experiments provide somewhat contradictory evidence, indicating that this is a subtle and complex issue that deserves more research.

## 1 Introduction

According to IPCC [Trisos et al., 2022] (cf. also [OECD, 2009]), despite being among the least contributors to global climate change emissions, African countries have already suffered significant losses and damages due to climate change. Moreover, they are disproportionately vulnerable to projected risks. In this context, *climate information*, which encompasses collecting, interpreting and communicating climate-related data, plays a critical role in assessing climate impacts and designing effective mitigation strategies. However, a key challenge lies in improving the accessibility and dissemination of climate information [WMO, 2006]. Enhancing the accessibility to relevant climate information is vital for empowering individuals [Hayhoe, 2018], and institutions, to make informed decisions, strengthen resilience to climate impacts, and actively engage in mitigation and adaptation efforts. With better climate information, we can support more effective climate action and contribute to building a sustainable and resilient future for Africa and the global community.

We begin framing the problem focusing on *global* climate information needs. We then zoom into those specific to Ghana and South Africa, as these two countries have very different climatic stresses and local ecological ecosystems. By conducting a preliminary quantitative and qualitative analysis of climate-related questions and answers, we seek to gain insights into the topical structure of information needs of individuals in these two countries and globally. We also conduct a preliminary study of answers generated by humans and a Large Language Model (LLM), ChatGPT [OpenAI, 2023], which highlights subtleties and challenges. We believe that such a complex problem can help drive AI research toward a better understanding of such models and inform future research and interventions to improve the accessibility and delivery of climate information.

---

\*Corresponding author: [massi@google.com](mailto:massi@google.com)

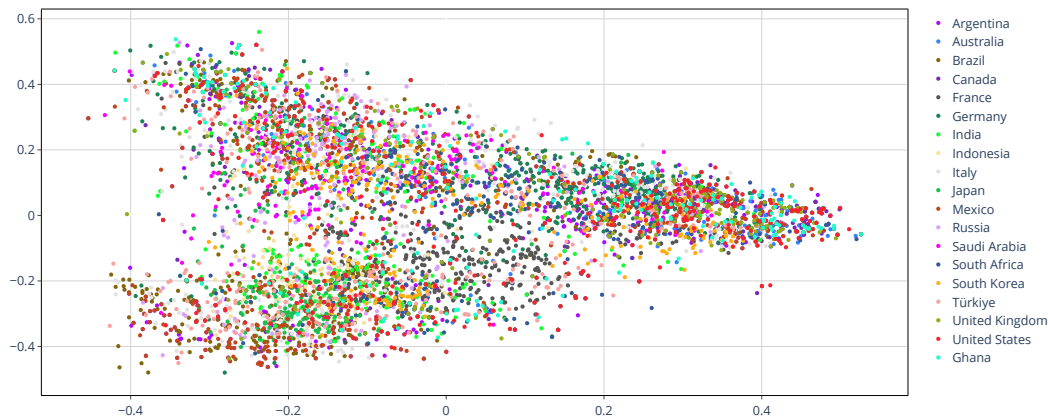


Figure 1: 2-dimensional PCA projection of the high-dimensional multilingual embeddings of all questions. The plot suggests the presence of a topical structure that is shared by all countries’ data.

## 2 Information Needs from Google Trends

To gain insights into the global climate-related information needs of individuals, we turn to Google Trends, a tool that provides data on the interest in specific search terms.<sup>2</sup>

### 2.1 The Epistemic Supercluster

We collect the top 500 questions, based on search volume, for the topics ‘Climate Change’ and ‘Global Warming’, from the period between 2020 and 2022, from each of the G20 countries, as well as Ghana.<sup>3</sup> We machine translate languages other than English for further examination. We filter out lengthy, off-topic, or unclear questions. To analyze the questions, we represent them in a shared embedding space, designed to capture semantic similarity across languages, by means of a language-agnostic embedding model trained on translation data [Feng et al., 2022]. To visualize the data, we apply Principal Component Analysis (PCA) [Pearson, 1901] to reduce the dimensionality.

The resulting projection, Figure 1, reveals a ‘flock of birds’-like structure where three broad topics emerge, shared among all countries. These topics concern basic questions about climate change (‘what is climate change?’), its causes (‘what are the causes of climate change?’), and lastly, questions regarding what can be done to address climate change (‘how to save the Earth?’). Perhaps not too surprisingly, this suggests that the dominant global climate information needs are epistemic. That is, most people want to understand what climate change is and how to deal with it. However, we remark that our study is only preliminary. In particular, we do not claim that the data represents any true underlying trends in any statistically accurate sense, nor that each country’s data is sampled in precisely equivalent ways.

### 2.2 Fine Grained Topics

To delve further into the questions’ semantic relations through the embeddings, we apply Uniform Manifold Approximation and Projection (UMAP) [McInnes et al., 2020]. UMAP is a flexible and scalable visualization technique which allows us to efficiently explore the data. We further apply k-means clustering to the resulting projection, to identify sub-clusters which we label with the corresponding central question. The result, see Figure 2, provides a better understanding of more specific topics. These include Basics (‘what is global warming?’), Causal/Consequences (‘how does climate change cause draughts’), How/Actions (‘how to scientifically solve the problem of climate change’), Controversial (‘global warming is a hoax’), Keyword-like (‘France global warming map’) and a catch-all category, Other. The cluster distribution can vary depending on the country.

<sup>2</sup><https://trends.google.com/trends/>.

<sup>3</sup>We do not have questions from China, as GTrends does not have data from this country. Thus, overall we consider Argentina, Australia, Brazil, Canada, France, Germany, Ghana, India, Indonesia, Italy, Japan, Mexico, Russia, Saudi Arabia, South Africa, South Korea, Türkiye, United Kingdom, and the United States.

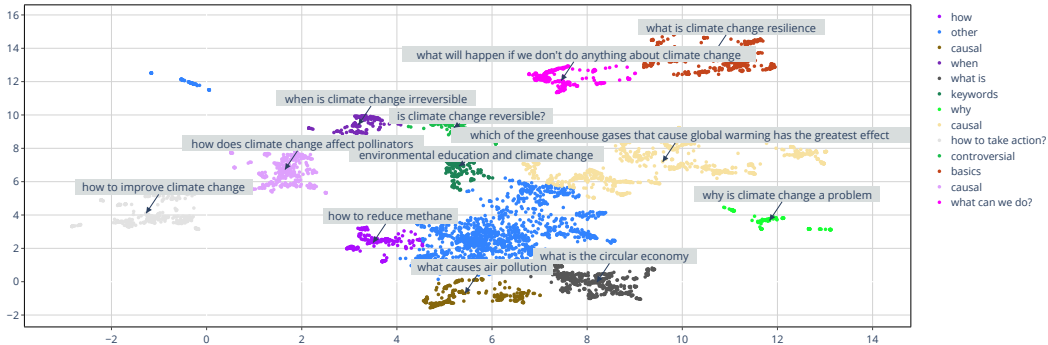


Figure 2: More fine-grained topics revealed by further analysis of the questions data.

### 3 A first look at Ghana and South Africa

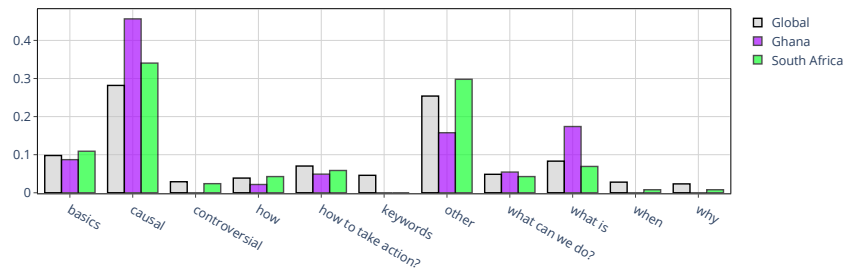


Figure 3: Topic distribution for Ghana (purple) and South Africa (green) compared to the global distribution (gray). Questions in the causal cluster are more frequent in Ghana and South Africa.

Within Africa, Ghana and South Africa lie across a wide span of latitude, around  $7.9^{\circ}$  N and  $30^{\circ}$  S, and are susceptible to different climatic threats that can have immediate consequences. For instance, sea-level rise and coastal economy are urgent concerns in Ghana, while in South Africa temperature extremes and less rainfall are top concerns for agriculture and related activities.

The cluster analysis for Ghana and South Africa, Figure 3, reveals common concerns about climate-related phenomena that are likely to impact the countries' economy. E.g. for Ghana, we find questions such as 'how does climate affect tourism in West Africa', 'how is climate change altering precipitation in Ghana' and 'how can climate change cause flooding and drought'. While for South Africa we find 'how does climate change affect agriculture', 'what is the relationship between climate change and the regularity of droughts in South Africa', 'how did the tropical cyclone Eloise impact the economy'.

These question reflect urgent information needs. Ghana's economy is highly vulnerable to climate change due to its heavy reliance on rainfall-dependent agriculture, with only a small fraction of irrigation potential utilized [Awuni et al., 2023]. Already today, smallholder farmers in Ghana's dryland farming systems use climate information to make critical farm decisions related to land [Antwi-Agyei et al., 2021b]. In north-eastern Ghana, access to and willingness to pay for climate information services are important factors for smallholder farmers [Antwi-Agyei et al., 2021a]. Most communities in Ghana, and Africa in general, broadly lack the requisite knowledge resources to inform adaptation choices [Manteaw et al., 2022].

Climate vulnerabilities in South Africa vary, as in Ghana, by region due to differences in geography, climate, and socioeconomic factors. For instance, the Western Cape and the Northern Cape provinces are prone to droughts and water scarcity, which can affect agriculture, water supply, and the overall ecosystem [Kapuka and Hlásny, 2021, Bannor et al., 2023]. This is particularly relevant and of concern to the wine industry in the Western and Northern Cape, which contributes over one percent of GDP [Bouri et al., 2018], and keeps South Africa among the top ten wine-producing nations in the world. The Eastern Cape and KwaZulu-Natal provinces are vulnerable to sea-level rise, coastal erosion, and extreme weather events such as storms and flooding [Schlosser et al., 2021], while

Gauteng province is vulnerable to the impacts of climate change on energy and infrastructure [Davis and Vincent, 2017]. Addressing regional vulnerabilities requires targeted adaptation and mitigation strategies, such as improving water management, promoting sustainable agriculture, investing in renewable energy technologies, and developing resilient infrastructure [Göktürk et al., 2023].

## 4 Human and AI answers to Climate Questions

Can AI, e.g., in the form of Large Language Models (LLMs) help answer climate-related concerns? What are its strengths and limitations? We ran the following experiments at the University of Ghana. We compiled a set of 93 climate questions relevant to Ghana (cf. Section B for more details). We then formed a group of 46 local experts from 13 sectors such as Forestry, Tourism, Energy, and Academia with (self-assessed) ‘moderate’ to ‘high’ knowledge of climate change. Divided into 4 groups, they wrote concise answers to each question. Afterwards, a different group from the one who wrote the answer assessed the answers’ quality as either ‘Low’ (4%), ‘Medium’ (56%), and ‘High’ (40%). At the same time, the groups assessed concise answers for the same question generated by ChatGPT. The result was that ChatGPT scored ‘High’ (99%) on all but one (‘Medium’) answers.

While ChatGPT was used as support for crafting the questions, the reason for this marked preference for the LLM-generated output, compared to responses given by humans with good knowledge on the subject, is still not clear. In what sense does an LLM give better answers than human experts? We hypothesize that the difference is, at least in part, explainable by surface properties such as style and narrative structure, possibly related to preference-based finetuning, RLHF [Ziegler et al., 2020].

Furthermore, a possible limitation of our experiment may be the setting in which the human experts work as a group in a limited time. This raises the question regarding how an LLM answer compares to an in-depth expert answer that was produced without such constraints. A systematic assessment is beyond the scope of this paper, but we analyze one example. We ask ChatGPT4 “How does climate change affect the agricultural sector in Ghana?” and receive a well-presented response that is not specific to Ghana (see Appendix C). We then identify a high-quality, human answer from a reputable source that delves deep into the effects of Climate Change on Ghana<sup>4</sup>. We extract the paragraph on ‘Agricultural Impacts’ as the human answer. To a human eye, the article’s answer is clearly superior. But we also ask ChatGPT which answer it prefers and why. This yields the following response:

“Considering that the question specifically refers to Ghana, [the human answer] is more appropriate as it gives specific details and impacts on the agricultural sector, including direct relevance to their main crop, cocoa, and how it might be affected.”

While ChatGPT might give answers that are deemed superior compared to humans’ in a time-constrained group setting, both humans and ChatGPT would prefer an expert answer that provides more relevant detail. Importantly, ChatGPT was able to recognize the better answer that it was not able to produce itself. More research is needed to improve our understanding of these phenomena.

## 5 Conclusion

This paper provides initial evidence of global and country-level patterns in epistemic demand for climate information. Moreover, the study shows that while LLMs can provide information deemed of high quality, their output can be sub-optimal; e.g., in terms of specificity. One challenge for AI research lies in obtaining high-quality answers from a system when it surpasses the collective capabilities of a group of experts in constrained settings. Additionally, evaluating the accuracy of the system’s content, even for experts, remains an open problem. This issue is being addressed in alignment and scalable oversight [Amodei et al., 2016]. Devoting more resources to climate information tasks could open up crucial opportunities for further advancement in this area, given the connection to science and science communication. We hope that climate information can grow its limited presence in AI datasets (cf. [Diggelmann et al., 2020]). We are particularly interested in datasets dedicated to African languages (cf. e.g., [Ogundepo et al., 2023]). Lastly, we highlight the main limitation of this work: We only provide (a first glimpse at) data from two African countries. The applicability of these ideas to a more representative sample of African countries and languages needs to be investigated, as well as the potential for translating findings into actionable strategies.

<sup>4</sup><https://www.climaterealityproject.org/blog/how-climate-crisis-impacting-ghana>

**Acknowledgements.** We would like to thank Vinodkumar Prabhakaran for his valuable feedback.

## References

- D. Amodei, C. Olah, J. Steinhardt, P. F. Christiano, J. Schulman, and D. Mané. Concrete problems in AI safety. *CoRR*, abs/1606.06565, 2016. URL <http://arxiv.org/abs/1606.06565>.
- P. Antwi-Agyei, K. Amanor, J. N. Hogarh, and A. J. Dougill. Predictors of access to and willingness to pay for climate information services in north-eastern Ghana: A gendered perspective. *Environmental Development*, 37:100580, 2021a.
- P. Antwi-Agyei, A. J. Dougill, and R. C. Abaidoo. Opportunities and barriers for using climate information for building resilient agricultural systems in Sudan savannah agro-ecological zone of north-eastern Ghana. *Climate services*, 22:100226, 2021b.
- S. Awuni, F. Adarkwah, B. D. Ofori, R. C. Purwestri, D. C. H. Bernal, and M. Hajek. Managing the challenges of climate change mitigation and adaptation strategies in Ghana. *Heliyon*, 2023.
- F. Bannor, I. H. Magambo, J. Mahabir, and J. M. Tshitaka. Interdependence between climate change and migration: Does agriculture, geography, and development level matter in sub-Saharan Africa? *South African Journal of Economics*, 2023.
- E. Bouri, R. Gupta, W.-K. Wong, and Z. Zhu. Is wine a good choice for investment? *Pacific-Basin Finance Journal*, 51:171–183, 2018. ISSN 0927-538X. doi: <https://doi.org/10.1016/j.pacfin.2018.07.002>. URL <https://www.sciencedirect.com/science/article/pii/S0927538X17305887>.
- C. L. Davis and K. Vincent. *Climate risk and vulnerability: A handbook for southern Africa*, 2017.
- T. Diggelmann, J. Boyd-Graber, J. Bulian, M. Ciaramita, and M. Leippold. CLIMATE-FEVER: A dataset for verification of real-world climate claims. In *Tackling Climate Change with Machine Learning workshop at NeurIPS 2020*, 2020.
- F. Feng, Y. Yang, D. Cer, N. Arivazhagan, and W. Wang. Language-agnostic BERT sentence embedding. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 2022. URL <https://aclanthology.org/2022.acl-long.62>.
- O. M. Göktürk, S. P. Sobolowski, M. H. Simon, Z. Zhang, and E. Jansen. Sensitivity of coastal southern African climate to changes in coastline position and associated land extent over the last glacial. *Quaternary Science Reviews*, 300:107893, 2023.
- K. Hayhoe. When facts are not enough. *Science*, 360(6392):943–943, 2018. doi: [10.1126/science.aau2565](https://doi.org/10.1126/science.aau2565). URL <https://www.science.org/doi/abs/10.1126/science.aau2565>.
- A. Kapuka and T. Hlásny. Climate change impacts on ecosystems and adaptation options in nine countries in southern Africa: What do we know? *Ecosphere*, 12(12):e03860, 2021.
- B. O. Manteaw, A.-B. Amoah, B. Ayittah, and K. B. Enu. Climate-informed decision-making in data-poor environments: Managing climate risk through citizen science networks. *Frontiers in Climate*, 4:835768, 2022.
- L. McInnes, J. Healy, and J. Melville. Umap: Uniform manifold approximation and projection for dimension reduction, 2020.
- OECD. Climate change and Africa. 2009. doi: [https://doi.org/https://doi.org/10.1787/gen\\_papers-v2009-art2-en](https://doi.org/https://doi.org/10.1787/gen_papers-v2009-art2-en). URL [https://www.oecd-ilibrary.org/content/paper/gen\\_papers-v2009-art2-en](https://www.oecd-ilibrary.org/content/paper/gen_papers-v2009-art2-en).
- O. Ogundepo, T. R. Gwadabe, C. E. Rivera, J. H. Clark, S. Ruder, D. I. Adelani, B. F. P. Dossou, A. A. DIOP, C. Sikasote, G. Hacheme, H. Buzaaba, I. Ezeani, R. Mabuya, S. Osei, C. Emezue, A. N. Kahira, S. H. Muhammad, A. Oladipo, A. T. Owodunni, A. L. Tonja, I. Shode, A. Asai, T. O. Ajayi, C. Siro, S. Arthur, M. Adeyemi, O. Ahia, A. Anuoluwapo, O. Awosan, C. Chukwunkeke, B. Opoku,

A. Ayodele, V. Otiende, C. Mwase, B. Sinkala, A. N. Rubungo, D. A. Ajisafe, E. F. Onwuegbuzia, H. Mbow, E. Niyomutabazi, E. Mukonde, F. I. Lawan, I. S. Ahmad, J. O. Alabi, M. Namukombo, M. Chinedu, M. Phiri, N. Putini, N. Mngoma, P. A. Amuok, R. N. Iro, and S. Adhiambo. AfriQA: Cross-lingual open-retrieval question answering for African languages, 2023.

OpenAI. Gpt-4 technical report, 2023.

K. F. Pearson. LIII. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine Series 1*, 2:559–572, 1901.

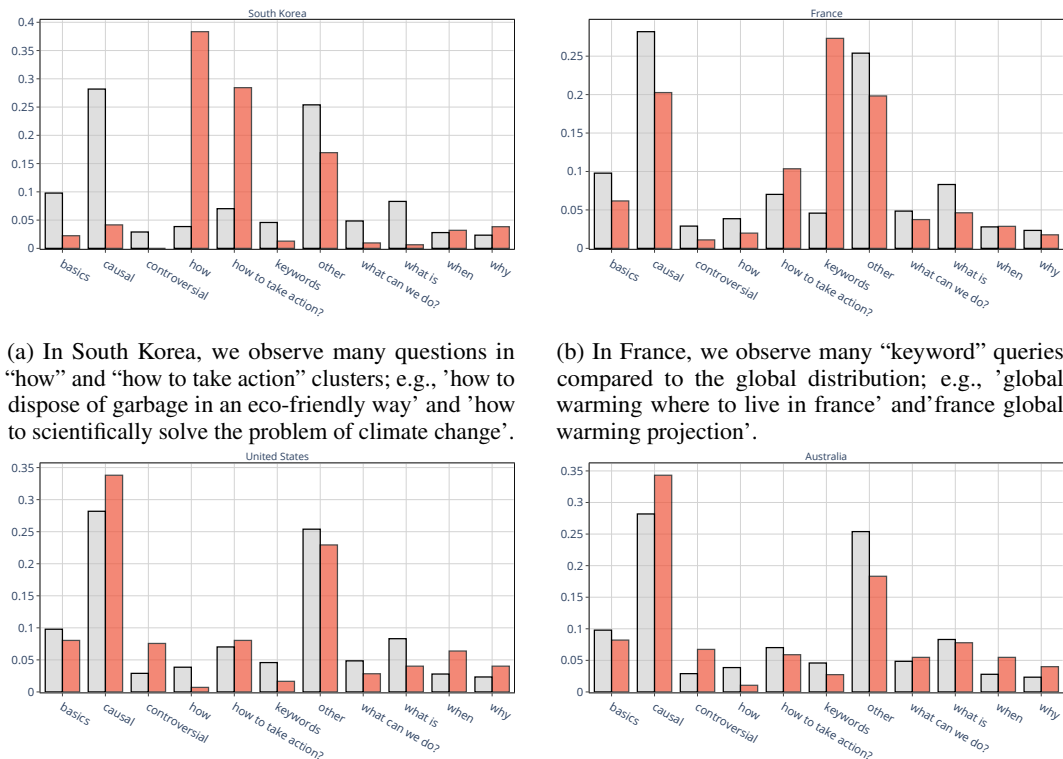
A. Schlosser, A. Sokolov, K. Strzepak, T. Thomas, X. Gao, and C. Arndt. The changing nature of hydroclimatic risks across South Africa. *Climatic Change*, 168:1–25, 2021.

C. Trisos, I. Adelekan, E. Totin, A. Ayanlade, J. Efitre, A. Gemedo, K. Kalaba, C. Lennard, C. Masao, Y. Mgaya, G. Ngaruiya, D. Olago, N. Simpson, and S. Zakieldean. *Africa*, pages 1285–1455. Cambridge University Press, Cambridge, UK and New York, USA, 2022. ISBN 9781009325844. doi: 10.1017/9781009325844.011.1286.

WMO. Climate information for development needs: an action plan for Africa - report and implementation strategy. 2006. URL [https://library.wmo.int/doc\\_num.php?explnum\\_id=3814](https://library.wmo.int/doc_num.php?explnum_id=3814).

D. M. Ziegler, N. Stiennon, J. Wu, T. B. Brown, A. Radford, D. Amodei, P. Christiano, and G. Irving. Fine-tuning language models from human preferences, 2020.

## A Appendix



(a) In South Korea, we observe many questions in “how” and “how to take action” clusters; e.g., ‘how to dispose of garbage in an eco-friendly way’ and ‘how to scientifically solve the problem of climate change’.

(b) In France, we observe many “keyword” queries compared to the global distribution; e.g., ‘global warming where to live in france’ and ‘france global warming projection’.

(c) In United States, we observe more questions in the “controversial” cluster than average; e.g., ‘climate change is fake’ and ‘global warming is a hoax’.

(d) In Australia we observe many “causal” queries such as ‘how does climate change affect {the great barrier reef, the daintree rainforest, food security}’.

Figure 4: Topic distributions for South Korea, France, United States and Australia.

## B Answer Generation/Assessment Experiment

The experiment explored 13 different sector/thematic areas related to climate information needs. These areas ranged from Climate Science to Health, Tourism, Academia, Aquaculture, Forestry, Agriculture, Waste Management, Transportation, Education, Water, Energy, and Energy Efficiency in Buildings. Each sector represents a unique aspect of climate information essential for understanding and addressing climate-related challenges. The questions covered the following topics.

toprule Sector/Thematic Area	Count
Climate Science	21
Health	11
Tourism	8
Academia	5
Aquaculture	5
Forestry	5
Agriculture	5
Waste Management	5
Transportation	5
Education	5
Water	5
Energy	5
Energy Efficiency in Buildings	5
Urban Development	4

While evaluating answers, the human experts were not revealed the source (human or ChatGPT) of the answer.

## C ChatGPT4 vs Expert Human example

The human example is taken from the website of the Climate Reality Project<sup>5</sup>.

The answer produced by ChatGPT4 is the following:

Climate change affects the agricultural sector in Ghana in several ways:

1. **Changes in Rainfall Patterns:** Climate change can lead to unpredictable rainfall patterns, which can disrupt the planting and harvesting schedules of farmers. This can lead to crop failures and food shortages.
2. **Increased Temperature:** Rising temperatures can lead to increased evaporation, reducing the amount of water available for irrigation and other agricultural uses. This can also lead to heat stress on crops, reducing yields.
3. **Increased Pest and Disease Incidence:** Changes in climate can create more favorable conditions for pests and diseases, which can damage crops and reduce yields.
4. **Soil Degradation:** Changes in rainfall and temperature can lead to increased soil erosion and degradation, reducing the fertility of the soil and making it harder for crops to grow.
5. **Reduced Crop Diversity:** Changes in climate can make it harder for certain crops to grow, reducing the diversity of crops that can be grown and potentially leading to food shortages.
6. **Impact on Livestock:** Climate change can also affect livestock, through heat stress, changes in feed quality and availability, and increased incidence of diseases.
7. **Economic Impact:** All these factors can lead to reduced agricultural productivity, which can have a significant economic impact on farmers and the wider economy. This can lead to increased poverty and food insecurity.

---

<sup>5</sup><https://www.climaterealityproject.org/blog/how-climate-crisis-impacting-ghana>