ClimaFactsKG: Towards an Interlinked Knowledge Graph of Scientific Evidence to Fight Climate Misinformation

Grégoire Burel[1], Harith Alani[1]

[1]Knowledge Media Institute, The Open University, Milton Keynes, UK

Abstract

Despite the overwhelming scientific evidence supporting the impact of humans on the environment, climate misinformation remains pervasive. This persistent spread of falsehoods is often achieved through the misrepresentation of scientific evidence and the promotion of pseudoscientific narratives that hinder effective climate action. To combat this issue, we introduce ClimaFactsKG, a knowledge graph that links common climate change denial narratives with scientific corrections. ClimaFactsKG currently consists of 252 common climate *myths* and the corresponding scientific counter-arguments. A key feature of ClimaFactsKG is its strategic integration with CimpleKG, one of the largest existing misinformation knowledge graphs. This connection allows the interlinking of scientific corrections with over 611 misinforming climate claims found in CimpleKG and significantly enhances the utility of ClimaFactsKG. By providing a structured and interlinked repository of climate change *myths* and their scientific rebuttals, ClimaFactsKG offers a valuable resource for researchers studying climate misinformation, fact-checkers seeking reliable counter-evidence, and educators aiming to improve climate literacy. This work represents a crucial step towards developing more robust, data-driven approaches to identify, analyse, and mitigate the effects of climate misinformation.

Keywords

Climate Misinformation, Knowledge Graph, Linked Data, Scientific Fact-Checking

1. Introduction

Although the scientific consensus about the role of human activity in altering the climate is overwhelming,[1] climate misinformation remains pervasive. The persistent spread of climate misinformation both online and offline cultivates scientific misunderstanding that hinders the understanding of climate change by the general public and the development of effective responses to the climate crisis.

In response to the rise of climate misinformation, several key strategies have emerged to combat its spread. These include fact-checking initiatives that directly debunk false claims, the development of climate scientific resources to organise and provide access to accurate information, and research-led communication strategies and tools that help identify and classify misinforming claims (see Section 2).

To deal with the continuously expanding amount of climate misinformation, fact-checking and research initiatives require easy access to scientific resources and historical climate misinformation. Although many websites offer scientific resources on climate misinformation, their human-centric design makes the content difficult for machines to access.

In this paper, we introduce ClimaFactsKG, a climate Knowledge Graph (KG) that connects climate misinformation *myths* with scientific corrective evidence. ClimaFactsKG relies on data extracted from the SkS website and connects it with the CARDS taxonomy [2] [1] of climate misinformation as well as CimpleKG [2], a large KG of fact-checks. By interlinking these three resources, ClimaFactsKG effectively bridges scientific corrections with real-world climate misinformation and facilitates access to corrective climate knowledge for a better understanding of the climate misinformation landscape.

At the time of writing,[3] ClimaFactsKG contains 252 common climate *myths* with their scientific

5th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment, Nov 2025, Nara, Japan

gregoire.burel@open.ac.uk (G. Burel); harith.alani@open.ac.uk (H. Alani)

0000-0003-0029-5219 (G. Burel); 0000-0003-2784-349X (H. Alani)

[1]NASA scientific consensus, https://science.nasa.gov/climate-change/scientific-consensus/.

[2]CARDS, https://cardsclimate.com.

[3]The statistics provided in the article are based on data collected on the 8th of July 2025.

correction. These *myths* are linked with 611 fact-checks from CimpleKG.

2. Related Work

Although there are some information sources designed to provide scientific knowledge related to climate misinformation, these resources tend to be available in a format that is primarily designed for human readability (i.e., website instead of an API or downloadable database) and come mostly from the fact-checking community. As a result, these resources cannot be easily integrated into automated fact-checking workflows and analysed easily. Skeptical Science[4] (SkS) is one of the most established sources of scientific knowledge for fighting common climate misinformation and *myths*. SkS is backed by multiple climate scientist.[5] and the website main purpose is to *"debunk misinformation that is harming our species' ability to deal with climate change caused by excessive anthropogenic greenhouse gas emissions"*.[6] This purpose is achieved by presenting peer-reviewed science in a concise form and explaining the techniques used for climate science denial. This takes the form of articles around specific climate change *myths*. Each article contains two boxes containing a summary of the misinforming claim (*"Climate Myths..."*) and the scientific correction (*"What the science says..."*). Detailed scientific explanations, including links to peer-reviewed research, can be found in the main content of each article, and some articles are provided in different languages with various levels of detail and complexity. The Climate Disinformation Database[7] (DeSmog) is also a relevant source of climate misinformation, but contrary to SkS, it provides information about individuals and organisations spreading climate misinformation rather than knowledge about common climate *myths*. A similar focus on individuals and organisations is also followed by Hot Air.[8] The CLIMATE-FEVER dataset [3] is perhaps the only accessible scientific source on climate misinformation. CLIMATE-FEVER is generated using English Wikipedia[9] data and relies on automatic methods to identify claims and scientific evidences. Due to its automatic nature, this approach may lead to inaccuracies when extracting scientific evidence.

Even though climate fact-checking data has been produced by fact-checking organisations (e.g., Climate Feedback[10], Euro Climate Check[11]), fact-checkers tend to only cover specific misinformation instance and do not provide scientific knowledge that can be easily reused outside the fact-check context. Although fact-checking KGs such as ClaimKG [4] and CimpleKG [2] exist, these resources do not explicitly identify climate misinformation and need to be separately classified using external knowledge sources or classification tools.

Classification tools and efforts for categorising and identifying climate misinformation have been created by researchers [1, 5, 6]. Most notably, the CARDS project [1] created both a taxonomy of climate misinformation and a climate claim classifier[12] that is based on data from the SkS website. The CARDS classifier is based on RoBERTa [7] and can classify the first two levels of the climate misinformation taxonomy except for the *"Climate is conspiracy*" category. A more recent version of the CARDS classifier, called Augmented CARDS [5], was developed for a similar purpose on top of DeBERTa [8], but contrary to the original CARDS classifier, it can also identify claims about the *"Climate is conspiracy"* category. Although this classifier is more targeted at content from X,[13], evaluation results show similar classification results between both classifiers when classifying the original CARDS data. The full description of the taxonomy categories can be found on the CARDS website.[14]. Climinator [6] relies on Retrieval-augmented generation (RAG) and climate corpora and focuses on identifying

[4]Skeptical Science, https://skepticalscience.com.
[5]What scientists are saying about SkS, https://skepticalscience.com/endorsements.shtml.
[6]About Skeptical Science, https://skepticalscience.com/about.shtml.
[7]Climate Disinformation Database (DeSmog), https://www.desmog.com/climate-disinformation-database/.
[8]Hot Air tool, https://www.tortoisemedia.com/hot-air-explore-tool.
[9]Wikipedia, https://en.wikipedia.org/
[10]Climate Feedback, https://climatefeedback.org.
[11]Euro Climate Check, https://euroclimatecheck.com/.
[12]CARDS: Computer-assisted recognition of (climate change) denial and skepticism, https://github.com/traviscoan/cards.
[13]X, https://x.com.
[14]Cards Climate, https://cardsclimate.com.

misinforming climate claims rather than classifying climate misinformation. Such a model is useful when climate claims have already been identified and share similarities with other automatic fact-checking approaches [9]. Approaches for generating corrections have also been investigated [10]. These approaches necessitate scientific corpora to create corrections to be effective.

3. The Climate Facts Knowledge Graph (ClimaFactsKG)

The Climate Facts Knowledge Graph (ClimaFactsKG) is a KG that integrates data from the SkS website, the CARDS taxonomy and the CimpleKG fact-checks. The following sections detail the methodology and steps involved in constructing ClimaFactsKG.

3.1. Skeptical Science, CimpleKG and CARDS

Finding a good source of information for scientific climate knowledge that can be easily associated with common climate misinformation is key to empowering climate fact-checks and research on climate misinformation. For constructing ClimaFactKG, we decided to rely on the data found on the SkS website as it provides common climate misinforming claims with short countering scientific corrections and CimpleKG [2] as it is a continuously updated KG of fact-checks (Section 2). Currently, CimpleKG contains more than 265,127 fact-checked claims from 94 organisations in 29 languages. In this paper, we use data from both SkS and CimpleKG to provide an interlinked KG that connects scientific climate knowledge for SkS to fact-checks. This integration allows for the analysis of both historical and current fact-checks using the climate *myths* identified by the SkS website.

A good categorisation of climate *myths* and corrections is needed for the easy identification of potential corrections of climate claims and the thematic analysis of climate misinformation. Although SkS provides a taxonomy of *myths*,[15] its structure is not always synchronised with the claims found on the website. As previously highlighted in Section 2, the CARDS project [1] provides a taxonomy of climate misinformation and climate claim classifiers that can be used to classify climate misinformation. Therefore, in this paper, we use the CARDS taxonomy to connect SkS data to CimpleKG resources.

3.2. The ClimaFactsKG Data Model

In this paper, we rely on the Schema.org vocabulary[16] (denoted with the `sc` prefix in the rest of this document) and the `sc:Claim` and `sc:ClaimReview` concepts for representing the claims and scientific corrections found in SkS since they align well with the structure of the SkS website data. Another reason for using the Schema.org ontology is that CimpleKG is also based `sc:ClaimReview`. This makes the integration between ClimaFactsKG and CimpleKG simpler as both KG share the same data model. `sc:ClaimReview` is also the standard used by fact-checking organisations to represent and publish their fact-checks. This data is generally embedded in fact-checking articles as JSON-LD.[17] By following the `sc:ClaimReview` format, ClimaFactsKG makes it possible to easily relate fact-check and claims to scientific evidence from the SkS.

The claims (`sc:Claim`) and corrections (`sc:ClaimReview`) found in the SkS pages are linked through the `sc:itemReview` relation. As SkS only provides corrections to misinforming claims, the rating of the claims is fixed to 0 via the `sc:ratingValue` property of the `sc:Rating`. The `sc:Rating` are then associated with the `sc:ClaimReview` using the `sc:reviewRating` relation. To ensure that these claims are never perceived as trustworthy, the `sc:worstRating` property is set to 0 and the `sc:bestRating` property is set to 1.

When multiple languages are available, language attributes are added to the textual attributes. If more than one version of the SkS article is found (i.e. complexity/explanation level), only the simplest

[15] Climate Myths sorted by taxonomy, https://skepticalscience.com/argument.php?f=taxonomy.
[16] Schema.org, https://schema.org.
[17] JSON-LD, https://json-ld.org.

available version of the article is made available as `sc:ClaimReview`. In the future, we will look at making the other versions available in ClimaFactsKG.

The content and metadata of the article associated with the climate *myths* (e.g., keywords, publication date, authorship, license and related articles) are linked to the `sc:ClaimReview` using relevant properties and relationships (e.g., `sc:keywords`, `sc:description`, `sc:abstract`) from Schema.org. The main mappings between the SkS content on `sc:ClaimReview` can be found in Table 1.

Table 1
Main Skeptical Science (SkS) *myths* mappings to the Schema.org ontology (`sc:ClaimReview`).

SkS Article Section		Mapping	Example Text from SkS Article
URL	→	`sc:ClaimReview` / `sc:url`	https://skepticalscience.com/global-cooling.htm.
What the science says...	→	`sc:reviewRating` / `sc:Rating` / `sc:ratingExplanation`	*All the indicators show that global warming is still happening.*
At a glance	→	`sc:abstract`	*Earth's surface, oceans and (...).*
Climate Myth...	→	`sc:claimReviewed` / `sc:Claim` / `sc:text`	*It's cooling "In fact global warming has stopped and a cooling is beginning (...).*
Last updated on (...)	→	`sc:dateCreated`	*4 June 2024.*
by (...)	→	`sc:author` / `sc:Person`	*John Mason.*
`<meta name="description"/>`	→	`sc:description`	*Empirical measurements of (...).*
`<meta name="keywords"/>`	→	`sc:keywords`	*global warming, skeptics, skepticism (...).*
`<title/>`	→	`sc:name`	*Global cooling - Is global warming still happening?*
Main content	→	`sc:reviewBody`, `sc:text`	*Earth's surface, oceans and atmosphere are all warming due to (...).*
Related Argument	→	`seeAlso`	https://skepticalscience.com/global-cooling-january-2007-to-january-2008.htm
source: (...)	→	`sc:citation`	https://wattsupwiththat.wordpress.com/2008/02/19/january-2008-4-sources-say-globally-cooler-in-the-past-12-months/

The connection to CimpleKG is achieved via the CARDS taxonomy, so the fact-checks found in CimpleKG can be easily linked to SkS concepts and scientific corrections. The CARDS taxonomy is mapped using the Simple Knowledge Organization System (SKOS) [11] as it is well-suited for representing knowledge hierarchies like taxonomies. The mapped SKOS concepts are then attached to ClimaFactsKG `sc:ClaimReview` using the `sc:about` property and inverse property `sc:subjectOf`. These connections can be queried to connect ClimaFactsKG `sc:ClaimReview` to CimpleKG `sc:ClaimReview`.

The ClimaFactsKG data and code are made available on GitHub[18] and the ClimFactsKG namespace is: https://purl.net/climatesense/climafactskg/ns#.

3.3. Data Collection, Annotation and Integration with CimpleKG and CARDS

As previously discussed, ClimaFactsKG is built on top of the climate *myths* articles found on the SkS website. For collecting the data, we created a web crawler that extracts the content of the SkS articles in a structured JSON format and then converts it to the `sc:ClaimReview` format used by ClimaFactsKG (Section 3.2). Although we collect all the available complexity levels for each article, we currently only map the lowest available level of each climate *myth* (typically this is the *Basic* level, but some articles have only the *Intermediate* complexity level). Although the crawler is currently manually executed, it can be updated to periodically update the scientific climate claims data found on SkS when new

[18]ClimaFactsKG code repository, https://github.com/climatesense-project/climafacts-kg.

claims are added to the website. The code of the crawler can be found on the GitHub repository of ClimaFactsKG.

After collecting the data, we use the Augmented CARDS classifier [5] for associating climate misinformation categories to the collected SkS data. We use this classifier rather than the original CARDS classifier [1] as it can identify the *"Climate is conspiracy"* category. The classification is performed using the content of the `sc:Claim` found in ClimaFactsKG. This classification is performed only on the English version of the SkS data since the Augmented CARDS classifier was trained on English data. However, since the different language versions of the same `sc:ClaimReview` are connected via the `sameAs` relation, it is easy to retrieve the CARDS classification for different languages.

To connect the SkS *myths* with existing fact-checks, we first identify the `sc:Claim` from CimpleKG that are related to the climate using the Augmented CARDS classifier. The CimpleKG `sc:ClaimReview` to the corresponding `sc:ClaimReview` in ClimaFactsKG are then connected through the CARDS taxonomy. To create these links, we query the CimpleKG Sparql endpoint using the query described in Listing 1. We only focus on the English `sc:ClaimReview` due to the language limitation of the Augmented CARDS classifier. Since sometimes the text of `sc:Claim` in CimpleKG is not in English despite the `sc:inLanguage` property, we use the *langdetect* library[19] to make sure that we only link to English claims.

```
PREFIX schema: <http://schema.org/>
PREFIX cimple: <http://data.cimple.eu/ontology#>
SELECT DISTINCT ?claimrev ?date_published ?text
WHERE {
   ?claimrev a schema:ClaimReview;
      schema:inLanguage "English";
      schema:datePublished ?date_published ;
      schema:itemReviewed ?claim .
   ?claim schema:text ?text .
}
ORDER BY DESC(?date_published)
```

Listing 1: SPARQL Query used for identifying CimpleKG claims that may be connected to ClimaFactsKG.

After collecting the SkS data and linking it to the CimpleKG `sc:ClaimReview` using the CARDS taxonomy, we identify 611 climate claims in CimpleKG that can be linked to the ClimaFactsKG. Currently, ClimaFactsKG consists of 252 base climate claims and scientific corrections that are available in up to 26 languages (not all 252 base claims are available in all the 26 languages).

4. Maintenance, Limitations and Future Work

To make it easy to use and integrate ClimaFactsKG into various research and fact-checking workflows, we decided to base the design of the KG on the Schema.org ontology and `sc:ClaimReview`. Although this simplifies the integration of the data with existing KGs such as CimpleKG, we identified some limitations of Schema.org in representing more complex content. For example, even though we collect all the versions of each SkS climate *myth*, we currently only represent the simplest available version of each claim, as `sc:ClaimReview` does not provide mechanisms for representing complexity variations of the same content. In the future, we plan to expand the Schema.org ontology so that the various versions of the content can be accessed through ClimaFactsKG.

Although new SkS *myths* are not added often, currently, the crawling SkS is not fully automated. Similarly, the mappings between CimpleKG and ClimaFactsKG are not yet automated. This means

[19]Langdetect library, https://github.com/Mimino666/langdetect.

that new fact-checks from CimpleKG are not mapped automatically as a new version of CimpleKG is released. We are currently working on automating such a task so ClimaFactsKG remains synchronised with CimpleKG.

Two other areas of improvement are the validation of the claim classification using the CARDS taxonomy and the extraction of scientific publications from the content of the SkS articles. Future work should formally evaluate the quality of the CARDS classification on the SkS data and the CimpleKG claims, as the number of retrieved CimpleKG claims appears to be low compared to the amount of claims available in the KG. Additional matching methods should also be considered, such as identifying similar claims between ClimaFactsKG and CimpleKG using embedding similarity or the creation of a new CARDS classifier based on new manual annotations. The future version of ClimaFactsKG will include the links to the scientific publications associated to the *myth* correction. This will provide additional means for querying the data via scholarly KGs and identifying additional relevant scientific resources that can be used for fighting and understanding climate misinformation.

Concerning the maintainability and long-term availability of ClimaFactsKG, the code and KG snapshots are publicly available on the KG repository. We also created a persistent URL for the data releases so that the data will always be accessible even if the SkS repository changes. Although we currently do not host a SPARQL endpoint, an endpoint can be easily created using the code found in the repository. Additional information about ClimaFactsKG availability can be found in the *Resource Availability Statement* Section of this article.

5. Conclusions

We introduced the ClimaFactsKG semantic resource, a KG that connects scientific climate evidence extracted from the SkS website with fact-checks from CimpleKG [2] using the CARDS climate taxonomy [1]. By linking common scientific corrections about the climate with common misinformation *myths*, ClimaFactsKG provide an important source of information to study climate misinformation. For its initial release, ClimaFactsKG contains 252 *myths* that are translated into up to 26 languages. These *myths* and scientific corrections are linked to 611 English fact-checks from CimpleKG. For future work, we are currently working on various improvements of ClimaFactsKG, such as the extraction of scholarly articles from SkS and automating ClimaFactsKG updates.

Acknowledgments

This work was supported by the European CHIST-ERA program within the ClimateSense project (Grant ID ANR-24-CHR4-0002, EPSRC EP/Z003504/1).

Resource Availability Statement

In addition to the URLs previously mentioned in the paper, ClimaFactsKG can be resolved using the following two persistent URLs: 1) the source code and the releases are available at: https://purl.net/climatesense/climafactskg; 2) the KG is available using the following URL and namespace: https://purl.net/climatesense/climafactskg/ns#. The data collected to create ClimaFactsKG is licensed under a Creative Commons Attribution 4.0 International License (CC-BY 4.0).[20] This license is the most recent version of the license used by the Skeptical Science website (Creative Commons Attribution 3.0 International (CC-BY 3.0)).

[20] CC-BY 4.0, https://creativecommons.org/licenses/by/4.0/.

Declaration on Generative AI

During the preparation of this work, the authors used Google Gemini in order to: *Grammar, spelling check and rephrase sentences.* After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] T. G. Coan, C. Boussalis, J. Cook, M. O. Nanko, Computer-assisted classification of contrarian claims about climate change, Sci. Rep. 11 (2021) 22320.

[2] G. Burel, M. Mensio, Y. Peskine, R. Troncy, P. Papotti, H. Alani, Cimplekg: A continuously updated knowledge graph on misinformation, factors and fact-checks, in: G. Demartini, K. Hose, M. Acosta, M. Palmonari, G. Cheng, H. Skaf-Molli, N. Ferranti, D. Hernández, A. Hogan (Eds.), The Semantic Web – ISWC 2024, Springer Nature Switzerland, Cham, 2025, pp. 97–114.

[3] J. Bulian, J. Boyd-Graber, M. Leippold, M. Ciaramita, T. Diggelmann, Climate-fever: A dataset for verification of real-world climate claims, in: NeurIPS 2020 Workshop on Tackling Climate Change with Machine Learning, 2020.

[4] A. Tchechmedjiev, P. Fafalios, K. Boland, M. Gasquet, M. Zloch, B. Zapilko, S. Dietze, K. Todorov, Claimskg: A knowledge graph of fact-checked claims, in: 18th International Semantic Web Conference (ISWC), Springer, 2019, pp. 309–324.

[5] C. Rojas, F. Algra-Maschio, M. Andrejevic, T. Coan, J. Cook, Y.-F. Li, Hierarchical machine learning models can identify stimuli of climate change misinformation on social media, Commun. Earth Environ. 5 (2024).

[6] M. Leippold, S. A. Vaghefi, D. Stammbach, V. Muccione, J. Bingler, J. Ni, C. C. Senni, T. Wekhof, T. Schimanski, G. Gostlow, T. Yu, J. Luterbacher, C. Huggel, Automated fact-checking of climate claims with large language models, npj Climate Action 4 (2025). URL: http://dx.doi.org/10.1038/s44168-025-00215-8. doi:10.1038/s44168-025-00215-8.

[7] Z. Liu, W. Lin, Y. Shi, J. Zhao, A robustly optimized bert pre-training approach with post-training, in: Chinese Computational Linguistics: 20th China National Conference, CCL 2021, Hohhot, China, August 13–15, 2021, Proceedings, Springer-Verlag, Berlin, Heidelberg, 2021, p. 471–484. doi:10.1007/978-3-030-84186-7_31.

[8] P. He, X. Liu, J. Gao, W. Chen, Deberta: Decoding-enhanced bert with disentangled attention, arXiv preprint arXiv:2006.03654 (2020).

[9] Z. Guo, M. Schlichtkrull, A. Vlachos, A survey on automated fact-checking, Transactions of the Association for Computational Linguistics 10 (2022) 178–206. doi:10.1162/tacl_a_00454.

[10] F. Zanartu, Y. Otmakhova, J. Cook, L. Frermann, Generative debunking of climate misinformation, in: D. Stammbach, J. Ni, T. Schimanski, K. Dutia, A. Singh, J. Bingler, C. Christiaen, N. Kushwaha, V. Muccione, S. A. Vaghefi, M. Leippold (Eds.), Proceedings of the 1st Workshop on Natural Language Processing Meets Climate Change (ClimateNLP 2024), Association for Computational Linguistics, Bangkok, Thailand, 2024, pp. 46–62. doi:10.18653/v1/2024.climatenlp-1.4.

[11] A. Miles, S. Bechhofer, Skos simple knowledge organization system reference (2009).