

RESEARCH ARTICLE



WILEY

Reading between the lines: The intersection of research attention and sustainable development goals

Christoph Funk¹ | Elena Tönjes² | Ramona Teuber^{1,3} | Lutz Breuer^{1,4}

¹Center for International Development and Environmental Research (ZEU), Justus Liebig University Giessen, Giessen, Germany

²Faculty of Economics and Business Studies, Department of Statistics and Econometrics, Justus Liebig University Giessen, Giessen, Germany

³Justus Liebig University Giessen, Institute for Agricultural Policy and Market Research, Giessen, Germany

⁴Institute for Landscape Ecology and Resources Management (ILR), Research Center for Bio Systems, Land Use and Nutrition (iFZ), Justus Liebig University Giessen, Giessen, Germany

Correspondence

Christoph Funk, Center for International Development and Environmental Research (ZEU), Justus Liebig University Giessen, Senckenbergstrasse 3, Giessen 35390, Germany.
Email: christoph.funk@wi.jug.de

Funding information

Deutscher Akademischer Austauschdienst, Grant/Award Number: 57526248

Abstract

In September 2015, the United Nations (UN) adopted 17 Sustainable Development Goals (SDGs) to transform our world by 2030. The scientific discourse around these SDGs has expanded rapidly since then, highlighting the need for efficient analysis of the large amount of textual data using Natural Language Processing. Our research addresses this need by employing a zero-shot text classification for SDG-related scientific articles, which allows for a thorough examination of scholarly discourse and the relationship between research attention and SDG achievement. We introduce the Research Attention Index (RAI), a novel metric that quantifies the research attention each SDG receives within a specific country. Our study contributes to the existing literature by providing a holistic view of global research attention to the SDGs. It also demonstrates the effectiveness of zero-shot text classification for large-scale textual labeling, and underlines the relevance of abstract analysis in understanding SDG-related discourse. Moreover, we examine the (non)-linear relationship between the RAI and SDG achievement across countries. Our results indicate considerable variations in the scientific discourse across countries worldwide and reveal a complex, non-linear relationship between research attention and progress towards achieving the SDGs. This underscores the importance of understanding the dynamics between research attention and sustainable development outcomes.

KEYWORDS

natural language processing, research attention index, SDG, sustainable development goals

1 | INTRODUCTION

In September 2015, the United Nations (UN) launched an ambitious agenda aimed at transforming the world by 2030. To tackle urgent development issues in areas like gender equality, infrastructure, environment, and education, this plan emphasizes global collaboration among all nations. At the heart of this transformative 2030 Agenda are 17 Sustainable Development Goals (SDGs), consisting of 169 inter-related targets and monitored by 231 indicators, which every country

worldwide is expected to achieve (Pedercini et al., 2019; United Nations, 2016b).

The scientific discourse around these SDGs has been expanding rapidly. While there is existing literature on the 2030 Agenda and the 17 SDGs established by the UN, this topic will be explored extensively in the ‘Literature Review’ section of this paper. It is important to note, however, that much of the current literature tends to focus on small text samples or relies on predefined numerical indicators set by the UN (Fuso Nerini et al., 2019; Pedercini et al., 2019).

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. *Sustainable Development* published by ERP Environment and John Wiley & Sons Ltd.

Our study aims to contribute to this emerging field by applying state-of-the-art Natural Language Processing (NLP) techniques, specifically using pre-trained language models like Bidirectional Encoding Representations from Transformers (BERT) (Devlin et al., 2018), to a large corpus of SDG-related scientific articles. In doing so, we seek to uncover the underlying structure of the scientific discourse on the SDGs and to explore the relationship between research attention and SDG achievement.

Using a language model, we apply zero-shot text classification, a powerful NLP technique that can automatically classify texts into predefined categories without requiring any training data for those categories. This technique is particularly advantageous in the context of SDGs, where the classification categories (i.e., the 17 SDGs) are well-defined but labeled training data may be scarce or unavailable. Thus, we use zero-shot text classification to assign each scientific article in our dataset to one or more SDGs based on its content. The use of zero-shot text classification highlights the potential of NLP techniques to augment traditional methods in SDG research. By leveraging the capabilities of advanced language models, we can extract rich semantic features from text and conduct large-scale analysis that would be challenging, if not impossible, with manual methods. This method significantly differs from traditional text classification techniques in several key aspects. Primarily, it eliminates the need for a training phase, as it relies solely on the intrinsic semantics of the text and predefined labels. This approach inherently reduces biases commonly associated with training data and eliminates the costs and efforts involved in labeling such data. Moreover, its flexibility allows policymakers to easily tailor the model to suit specific requirements, ensuring that the classification process aligns closely with their evolving needs.

To better understand this relationship, we introduce a new Research Attention Index (RAI), to provide a quantitative measure of research attention dedicated to each SDG within a country, and thus providing a metric that can complement traditional indicators in capturing the dynamics of SDG-focused scientific discourse. By quantifying research attention, the RAI offers a novel perspective to understand how global research attention is distributed across different SDGs and countries. It enables us to explore whether there is a correlation between the level of research attention and the progress made in SDG achievement, providing insights that could guide research funding and policy decisions. Furthermore, it allows us to determine whether the distribution of research attention aligns with global development priorities, thereby identifying potential gaps or imbalances.

Our research contributions are four-fold. First, we provide a comprehensive overview of global research attention on SDGs. Second, we demonstrate how zero-shot text classification can efficiently label extensive textual information, providing policymakers with information beyond the typical UN indicators in an efficient manner. Third, we introduce the novel RAI to quantify each SDG's research attention per country, and we analyze its (non)-linear relationship with SDG achievement. Fourth, we showcase that abstracts are sufficient for capturing the most relevant SDG-related information in scientific articles, thereby saving computational resources.

The remainder of this paper is organized as follows: The section 'Literature Review' provides an overview of the relevant literature related to our work. In the 'Data' section, we describe our dataset, providing a descriptive analysis that emphasizes key characteristics of our texts, such as the word counts related to each SDG, and the geographical distribution of our research articles. This section is followed by the 'Methodology' section, which explains the techniques we employed, including zero-shot classification, and presents an overview of the labels used for our analysis. The 'Abstracts versus Full Texts' section provides a comprehensive comparison between the use of abstracts versus full texts in the context of our RAI. In the 'Results' section, we present our primary findings followed by a discussion in the 'Discussion' section. Concluding remarks can be found in the 'Conclusion' section.

2 | LITERATURE REVIEW

While the 2030 Agenda and its 17 SDGs, as launched by the UN, represent a global commitment to addressing pressing issues like gender equality, infrastructure, environment, and education, the scientific discourse on these topics is rapidly evolving. Existing literature often relies on small text samples or confines its assessment to predefined numerical indicators set by the UN (Fuso Nerini et al., 2019; Pedercini et al., 2019). For example, Bennich et al. (2020) manually assessed 70 peer-reviewed articles to provide an overview of SDG interactions in the scientific literature. Le Blanc (2015) conducted a network analysis to uncover relationships among SDGs based solely on the wording of the targets, while Bali Swain and Ranganathan (2021) performed a network analysis to reveal synergies and trade-offs exclusively based on the SDG targets. These small-scale or theory-based studies either capture only a subset of the scientific literature addressing the SDGs, potentially leading to a biased sample, or focus on theoretically assumed relationships among the SDGs.

Another segment of the existing literature consists of manual literature reviews that analyze and discuss a limited number of papers on the SDGs in general or on specific SDGs (Decouttere et al., 2021; Hackl, 2018). While valuable, these studies do not fully exploit the potential of large-scale text analysis. NLP offers a compelling alternative, capable of handling vast textual data and extracting meaningful patterns (Smith et al., 2021). A small but growing number of studies have employed NLP techniques to analyze SDG-related texts, uncovering valuable insights and potential collaboration opportunities. Sebestyén et al. (2020) used NLP to perform a network analysis of the SDGs in Voluntary National Reviews, uncovering informative word pairs to draw SDG connections between countries. The Voluntary National Reviews share experiences, successes, challenges, and lessons learned to accelerate the implementation of the 2030 Agenda for Sustainable Development (United Nations Department of Economic and Social Affairs, 2019). Findings from studies like this can help reveal the SDG areas where countries face similar challenges and where they might collaborate (Sebestyén et al., 2020). Another study

by Smith et al. (2021) used 85 UN progress reports on each SDG to measure the dependencies and interactions between SDGs and identify potential clusters. The authors also assessed whether their findings on SDG inter-dependencies based on the UN progress reports were reflected in scientific articles from the last two decades. Additionally, Chang et al. (2021) used cluster analysis and topic modeling on environmental education research journals from the Web of Science during 2011–2020 to analyze research topics. The techniques employed in Chang et al. (2021), such as encoding the text with term frequency-inverse document frequency and using a combination of Latent Dirichlet Allocation for topic modeling and K-means algorithm, are well-established concepts in NLP (Chang et al., 2021; Sebestyén et al., 2020; Smith et al., 2021).

Recent studies, such as Bellantuono et al. (2022), have begun integrating large language models like BERT for SDG analysis. They trained a BERT model for a multi-label classification task with the 17 SDGs as labels. Their training and analysis dataset consisted of Japanese texts from various sources. The primary aim was to harness the model's capabilities to predict connections among different SDGs using a dataset containing representative indicators. This approach aimed to uncover opportunities for uniting stakeholders interested in collaborating on SDG initiatives. Angin et al. (2022) also use a large language model, specifically utilizing an advanced RoBERTa model. This model excels in dissecting sustainability reports, identifying key sections related to SDGs. For development and validation, they employed the OSDG Community Dataset, which is extensively annotated with SDG-focused text by community members. Notably, this study leans towards a more technical and model-centric approach. It includes thorough comparative testing against both traditional machine learning and contemporary deep learning models, demonstrating the model's robustness and precision.

Despite the growing body of research employing NLP, there remains a need for more sophisticated, large-scale textual analysis that can capture the SDG discourse in the scientific community in depth. Our study takes a substantial step forward by employing the pre-trained language model BERT and the innovative technique of zero-shot text classification. By applying these techniques to an extensive corpus of SDG-related scientific articles, we aim to unveil the intricacies of SDG-focused scientific discourse, providing a detailed and comprehensive understanding. To address these gaps, our research proposes the following questions:

1. How can advanced NLP techniques, particularly pre-trained language models and zero-shot text classification, be effectively utilized to analyze and interpret large-scale textual data related to the SDGs?
2. What novel insights and patterns can these advanced NLP techniques reveal in the scientific discourse on SDGs that are not evident in smaller-scale or traditional studies?
3. How does the application of these techniques inform the understanding of the relationship between research attention (as quantified by our RAI) and the progress in achieving SDGs?

By answering these questions, our study aims to contribute significantly to the field. We intend to provide an in-depth analysis of global research attention on SDGs and explore its implications for global development strategies. This approach not only enhances the understanding of SDG-focused discourse but also informs future research and policy-making in the realm of sustainable development.

3 | DATA

In our study, we utilized the Web of Science (www.webofscience.com) to conduct a comprehensive search for articles related to 'Sustainable Development Goals'. As of May 07, 2023, we identified 17,098 articles that matched our criteria. The methodology for determining our dataset is outlined in Figure 1. Since our model is specifically designed to analyze English text data, we limited our dataset to articles written in English, which reduced the dataset by 513 articles.

To ensure relevance and alignment with the Agenda 2030, we focused on articles published between the years 2015 and 2022. This decision was made based on the assumption that articles written before 2015 might not consistently align with the SDGs. As a result, we excluded 1,406 further articles written prior to 2015. Furthermore, we chose not to incorporate articles from 2023 into our study, given that the corresponding SDG indicators for 2023 have not yet been released. The dataset utilized in our study comprises titles and abstracts provided by the Web of Science. These texts are machine-readable and available in English, even if the full articles are not, which made them easily accessible for our analysis.

However, we encountered a limitation in our dataset as 554 articles lacked abstracts in the Web of Science metadata. To maintain a concentrated focus on the SDGs, we included only articles that explicitly referenced the term in any form. We employed a regex function, detailed in the following section, to filter the articles accordingly. As a result, we excluded 1,487 articles from our dataset. This left us with a final count of 13,138 articles, which formed the basis for our analysis.

3.1 | Descriptive analysis

Descriptive analysis is crucial for understanding the distribution and prevalence of the 17 SDGs in our dataset. This enables us to provide an overview of the frequency distribution of each SDG, which can help researchers and policymakers identify the goals that are receiving more attention in the literature and potentially prioritize further research or interventions based on the analysis. To determine which of the 17 SDGs is most commonly mentioned in our dataset, we used regular expressions to count the occurrences of SDGs per article. We counted the acronym SDG and the trigram 'Sustainable Development Goal' as SDG, while the acronym SDGs and 'Sustainable Development Goals' were counted as SDGs. Furthermore, we included various combinations of each of the 17 SDGs, such as 'SDG1', 'SDG 1', 'SDG one', and combinations like 'SDG 1.1', 'SDG 1a', and so on, as SDG 1.

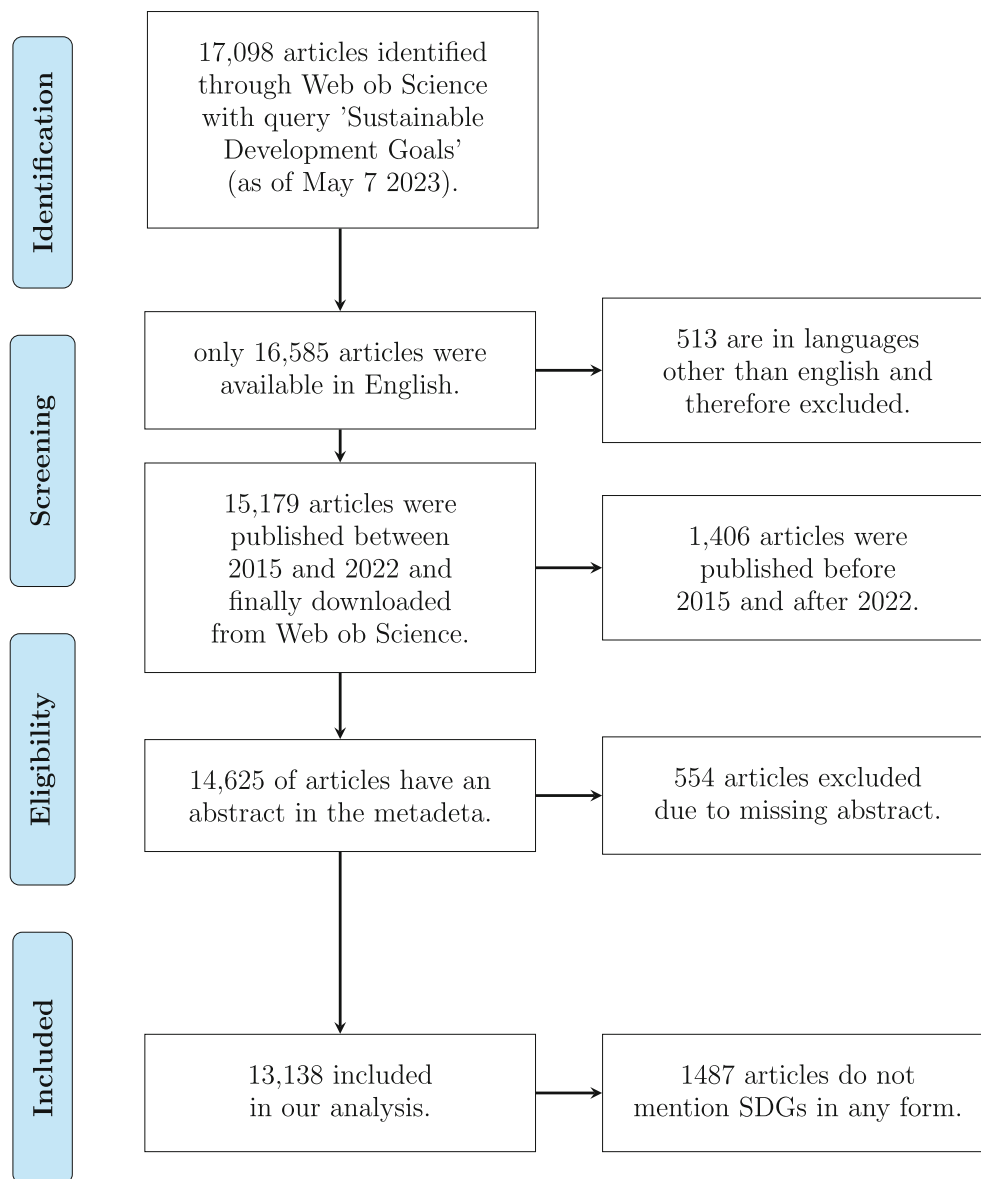


FIGURE 1 PRISMA flow diagram of SDG article data collection.

Table 1 presents a comprehensive overview of the total number of SDGs counted, the number of abstracts in which each specific SDG appears at least once, and the average number of mentions per abstract. As anticipated, the generic terms 'SDG' and 'SDGs' were most frequently used, appearing 36,366 times across 13,136 abstracts. On average, these terms were mentioned 2.77 times per abstract and title.

We noticed a significant variation in the frequency of mentions for each SDG throughout the dataset. SDG 3—*good health and well-being* was the most commonly cited goal, appearing in 395 abstracts with a total count of 564 mentions and an average of 1.43 mentions per abstract. SDG 6—*clean water a sanitation* followed as the second most frequently mentioned goal, featured in 288 abstracts with a total of 453 mentions and an average count of 1.57 mentions per abstract.

Conversely, SDG 16—*peace, justice, and strong institutions* was the least frequently cited goal in our dataset. It appeared in only 101 abstracts, with a total count of 150 mentions and an average of

1.49 mentions per abstract. This wide variation in the frequency of SDG mentions highlights the diverse range of research topics and focus areas within the context of sustainable development.

3.2 | SDG co-occurrence

By examining the co-occurrence of SDGs within articles, we can identify possible relationships or interdependencies among the goals. This can provide insights into how researchers and policymakers might address multiple SDGs simultaneously, leading to more efficient and effective interventions. Figure 2 illustrates the co-occurrence of SDGs in our dataset. The lines connecting the SDGs indicate the number of times two specific SDGs have been mentioned together in the same article. For clarity and readability, we have chosen to highlight only those connections that are present in 40 or more articles. This visualization aids in analyzing the interconnections among SDGs and

TABLE 1 Counted SDGs for all abstracts.

SDG	Short description	Total count	Number of abstracts	Average count per abstract
1	No poverty	249	184	1.35
2	Zero hunger	330	255	1.29
3	Good health and well-being	564	395	1.43
4	Quality education	301	212	1.42
5	Gender equality	206	160	1.29
6	Clean water and sanitation	453	288	1.57
7	Affordable and clean energy	340	257	1.32
8	Decent work and economic growth	233	193	1.21
9	Industry, innovation and infrastructure	137	115	1.19
10	Reduced inequalities	147	125	1.18
11	Sustainable cities and communities	339	238	1.42
12	Responsible consumption and production	258	211	1.22
13	Climate action	329	277	1.19
14	Life below water	213	135	1.58
15	Life on land	269	187	1.44
16	Peace, justice and strong institutions	150	101	1.49
17	Partnership for the goals	154	122	1.26
	SDG/SDGs	36,366	13,136	2.77

Note: Table presents a comprehensive overview of the total number of SDGs counted, the number of abstracts in which each specific SDG appears at least once, and the average number of mentions per abstract.

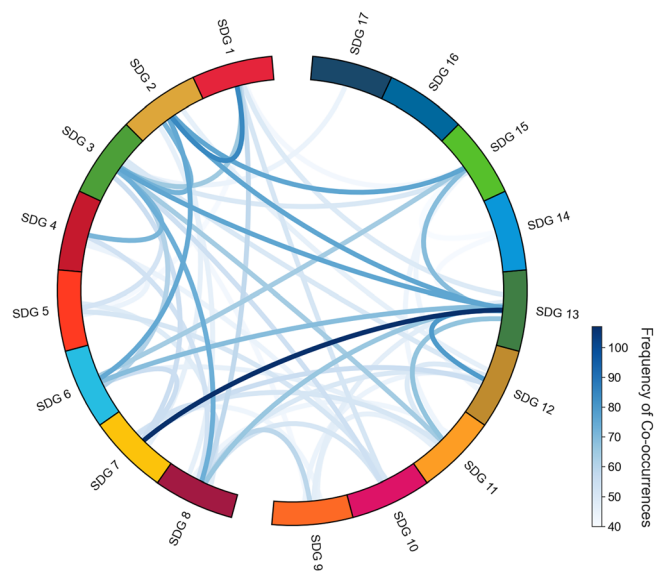


FIGURE 2 SDG co-occurrence plot. The lines connecting the SDGs represent the frequency at which two specific SDGs are mentioned together within the same abstract. The colorbar provides a visual representation of the strength of these relationships.

identifying the most frequently co-occurring ones in the research discourse. It is apparent that certain SDGs are cited together more often, suggesting robust interrelationships or intersecting research interests. On the contrary, SDGs with fewer co-occurrences may reflect a lesser degree of interconnectivity or a narrower focus of research concerning those specific objectives.

For example, SDG 7—*affordable and clean energy* and SDG 13—*climate action* have been mentioned together in 107 abstracts. This frequent co-occurrence is not unexpected, as these goals are intrinsically linked in addressing global environmental challenges. Both goals aim to tackle the negative impacts of human activities on the environment and promote sustainable development. The frequent co-occurrence of SDGs 7 and 13 in research articles reflects the interconnected nature of these goals and the common understanding that addressing energy and climate issues simultaneously is crucial for sustainable development.

Additionally, SDG 1—*no poverty* and SDG 2—*zero hunger* have been mentioned together the second most times with 84 articles, followed by SDGs 12 and 13 with 79 times. While SDG 1 aims to end poverty in all its forms, SDG 2 focuses on ending hunger and malnutrition. Hunger and malnutrition are directly linked to poverty, as people living in poverty often lack the resources to access or produce enough nutritious food for themselves and their families. These interconnected issues require integrated solutions to promote sustainable development and ensure food security for vulnerable populations.

Moreover, the co-occurrence of SDG 12—*responsible consumption and production* and SDG 13—*climate action* highlights their intrinsic connection through the understanding that responsible consumption and production are essential components of climate action. Sustainable production and consumption patterns help reduce greenhouse gas emissions, minimize waste, and conserve natural resources, ultimately contributing to climate change mitigation and adaptation.

In general, our analysis of the co-occurrence reveals that SDG 13—*climate action* appears to be a central theme, as it exhibits strong interlinkages with SDGs 2—*zero hunger*, 3—*good health and well-being*,

6—clean water and sanitation, 8—decent work and economic growth, 11—sustainable cities and communities, 12—responsible consumption and production, and 15—life on land. This central role of SDG 13 suggests that climate action and climate change as a global threat not only affects the environment but also has wide-ranging implications for human well-being, economic development, and social equity.

It is important to note that the number of mentions alone or the number of co-occurrences cannot conclusively determine the importance of individual SDGs or connections between them. For instance, an article may discuss actions required to combat climate change and its impacts without explicitly mentioning SDG 13—*climate action*. Therefore, this descriptive analysis should be seen as a foundation for more advanced text analysis techniques, such as zero-shot text classification discussed in our ‘Result’ section. These methods can offer a more nuanced understanding of the significance of each SDG in the dataset, as well as the broader context in which they are discussed, thereby providing a more comprehensive insight into the research landscape surrounding the SDGs.

3.3 | Geographical patterns

A further analysis provides valuable insights into the global scientific discourse surrounding the SDGs. By examining our dataset in terms of the countries mentioned, we can discern geographical patterns and relationships between articles and their associated countries. To achieve this, we use the ‘find countries’ function from the ‘country named entity recognition’ library in Python, which identifies countries in the combined text of each article’s title and abstract. The rationale behind this is to determine the main country or countries an article focuses on. If a country is mentioned in either the title or abstract, we assume that the article is centered on that country. In cases where an article discusses multiple countries simultaneously, it will be counted towards all mentioned countries.

We use Python along with the ‘pycountry’ library to extract country objects from the tuples and append the country codes (ISO 2 and ISO 3) to the ‘Articles’ DataFrame. The ‘country named entity recognition’ library is case-insensitive, ensuring accurate country detection regardless of text case variations. We manually verified the code’s effectiveness by randomly sampling 30 abstracts, and it performed well in all instances.

In the Appendix, Table A3 provides a summary of research article counts by ISO2 country code. The top five countries with the most articles are China (CN) with 825, India (IN) with 623, the United States (US) with 309, South Africa (ZA) with 265, and Brazil (BR) with 238. The table includes counts for numerous other countries with at least five research articles, illustrating the global scope of research attention.

4 | METHODOLOGY

4.1 | Zero-shot text classification

Text classification is a crucial NLP task for leveraging knowledge in texts. It involves categorizing sentences, paragraphs, or entire

documents into predefined labels. Deep neural models are commonly used for this task, but they require large amounts of labeled data for training. However, creating labeled data is expensive and time-consuming, and even human annotators can struggle to assign the most appropriate label to a text, despite fully understanding its semantics (Beltagy et al., 2019). Zero-shot classification is a technique that facilitates text classification without necessitating training data. Traditional classification models typically undergo training, during which they learn patterns from the text, enabling them to categorize new, previously unseen texts. However, zero-shot classification distinguishes itself by its ability to classify texts with minimal training. This is achieved by leveraging the inherent semantics of both the text and the labels, instead of depending solely on learned patterns. One approach to achieve this is by embedding sentences and labels into the same latent space and computing the distance between them. This approach depends exclusively on the semantics of the sequence and label for classification (Socher et al., 2013). The model we employ, developed by Davison (2020), utilizes a unique embedding process outlined as follows:

- Select the top K words, referred to as V , from the most frequently used words in a word2vec model’s vocabulary.
- Generate embeddings for these words through word2vec, represented as $\Phi_{\text{word}}(V)$.
- Similarly, produce embeddings for the same set of words using Sentence-BERT (S-BERT), indicated as $\Phi_{\text{sent}}(V)$.
- Develop a linear projection matrix Z , optimized by least-squares with $L2$ regularization, mapping $\Phi_{\text{sent}}(V)$ onto $\Phi_{\text{word}}(V)$.

Here, Z functions as an additional transformation layer, enhancing the S-BERT embeddings for both the sequence and the labels. However, our approach extends this concept by also interpreting sentences as labels, thereby allowing the mapping of entire sentences into the latent space. We adopt a variant of this technique that includes multi-natural language inference (MNLI) for advanced sentence classification. In this process, a given sequence is considered a premise, and the labels are formulated as hypotheses. These hypotheses are then evaluated against the premise to ascertain whether they entail, contradict, or are neutral to the premise. This method aligns with the architecture proposed by Yin et al. (2019). For instance, Davison (2020) provided an example where the sentence ‘A soccer game with multiple males playing’ is the premise, ‘Some men are playing a sport’ acts as the hypothesis, and the resulting label is ‘entailment’. This example aptly demonstrates how MNLI can be applied to classify sentences based on their inferred relationships. In previous work, Yin et al. (2019) transformed naturally classified pairs into binary categories, specifically ‘entailment’ versus ‘non-entailment’, to produce a binary outcome. They employed a pre-trained MNLI BERT model for zero-shot testing on this modified dataset. In our study, we have adapted the zero-shot classification model as developed by Davison (2020), which is available in the Hugging Face’s Transformers library (Wolf et al., 2020). This model leverages the Bidirectional and Auto-Regressive Transformers (BART) framework, which is used for

tokenizing both the input sequence and the label. The BART model, as described in Lewis et al. (2019), uniquely integrates a sequence-to-sequence translation framework, combining bidirectional encoding capabilities (similar to BERT) with unidirectional, left-to-right encoding (reminiscent of the GPT model). In context of sequence classification tasks like the MNLI problem addressed in our study, the BART model operates by feeding the same input into both its encoder and decoder components. The crucial aspect of this model is how it processes the input: the final hidden state token from the decoder is passed through a multi-class linear layer, producing a logit vector. Davison (2020) enhanced this model by applying a softmax function to the logit classification scores. This allows for the calculation of probabilities associated with ‘entailment’ and ‘contradiction’ while deliberately omitting the ‘neutral’ classification score. These probabilities are then interpreted as the likelihood that the labels accurately correspond to the sequences inputted into our system.

In summary, our zero-shot classification model relies on the pre-trained MNLI BART model, as developed by Davison (2020) and integrated in the transformer library (Wolf et al., 2020). This model's effectiveness lies in its ability to interpret and classify sequences by analyzing both the semantics of the input sequences and the labels fed into the model.

4.2 | Model validation

Zero-shot classification offers the advantage of not requiring a large labeled dataset for training the model. However, it also lacks a common validation procedure, such as splitting the labeled dataset into train and test sets to evaluate the model's performance. To address this, we validated our model and labels using the UN's annual progress reports (United Nations Economic and Social Council, 2016, 2017, 2018, 2019, 2020, 2021) for each SDG. These reports should contain similar semantics about SDG performance as scientific articles that write about SDGs. We used these reports as a proxy to validate our labels and model by labeling each sentence in the progress reports with the corresponding SDG. This resulted in a labeled dataset of 2,301 sentences without any labeling effort. We then applied the zero-shot classification model with our defined labels to all sentences and verified whether the probabilities matched the labels.

Figure 3 illustrates our results, and Table A4 in the Appendix provides a detailed overview of the labels we used. To test the performance of different label lengths, we used two labels for each SDG, a shorter and a longer version. Figure A3 in the Appendix presents the results for both versions separately, and Figure 3 shows the mean across the long and short labels for each SDG. Initially, we included

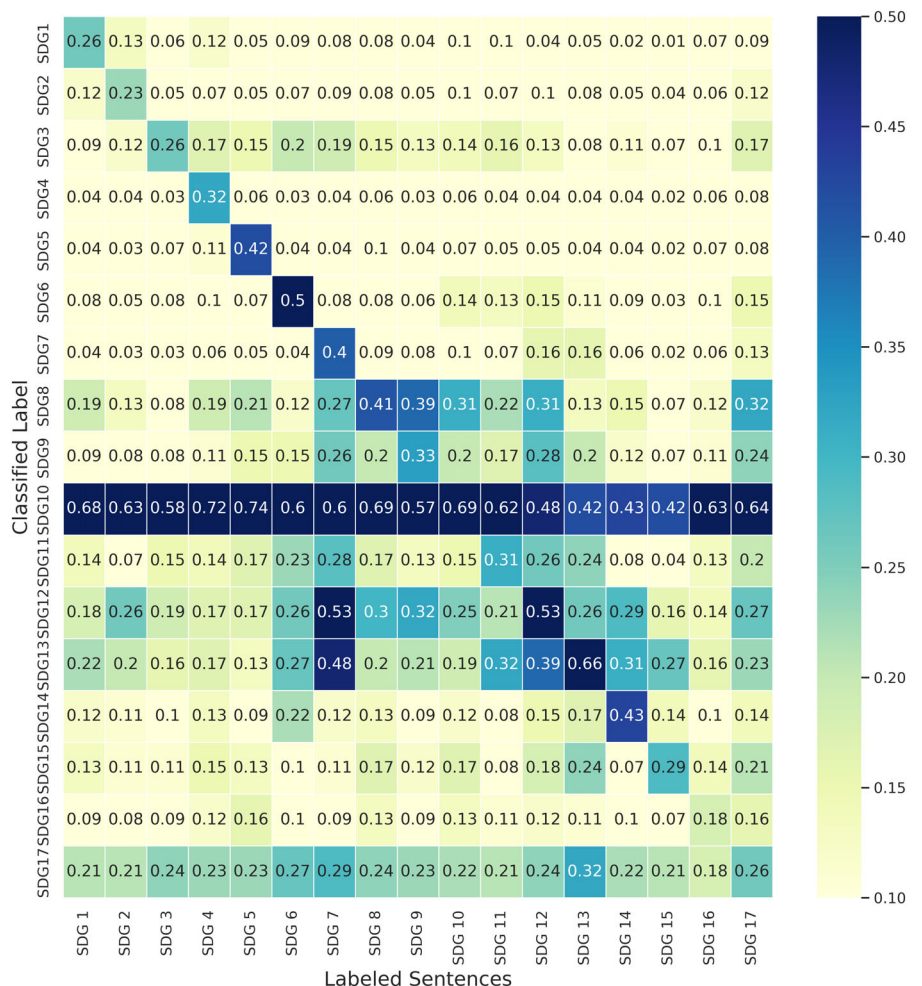


FIGURE 3 Model validation with mean values from short and long labels by comparing the sentences classified by the model on the y-axis with the labels classified by humans on the x-axis.

labels for all SDGs consisting of either the SDG number or the SDG number spelled out as a word (e.g., SDG 1 or SDG one), but we decided not to include these labels in our final analysis as they did not capture the semantic meaning of ‘SDG’ and performed poorly in our validation approach.

As we allow for multiple labels to be true, we have a probability distribution for each label over all pre-labeled sentences. The x-axis represents the sentences from the progress reports, grouped by the corresponding SDG and thus by the UN progress report label. The probability for a label should be highest for the sentence from the corresponding progress report, as the report should mostly contain sentences about the matching SDG. Therefore, for the version where we included the long and short label, both squares on the diagonal should have the darkest color, as a darker color indicates a higher probability. The same should hold for the average over both labels, where the squares on the diagonal should be darkest. In Figure A3 in the Appendix with the long and short versions, we can observe that the probabilities are typically highest on the diagonal, with one label usually performing slightly better than the other. Nonetheless, we opted to use the mean of both labels since we are only applying the model on a finite sample, and with a slightly different dataset, some results could be the opposite. Figure 3 presents a clearer picture of our results. In most cases, the diagonal is distinguished by the highest probability in the row. However, some labels have multiple high probabilities within their respective row. Notably, SDG 10 stands out as the probability of discussion is high for most SDG labels. This could be attributed to the fact that SDG 10 has numerous linkages to all the other UN SDGs (United Nations, 2016a). Moreover, “reduce inequality” can be applied to several other SDGs and can be framed as a general goal that needs to be achieved.

The validation heatmaps shown in Figure 3 reveal some interesting findings. Firstly, it is notable that sentences from all progress reports are related in some way to SDG 10, which aligns with the goals included in SDG 10 itself. Secondly, sentences related to SDG 7 are more likely to be classified as sentences related to SDG 12 and SDG 13, which can be explained by the linkages illustrated by the UN. The link between SDG 7 and SDG 13 and the link between SDG 13 and SDG 12 can also be seen in the linkages found in our co-occurrences in Figure 2. However, sentences on affordable energy seem to fit *climate action* and *responsible consumption and production*, but not vice versa. Furthermore, in section ‘SDG co-occurrence’ we find linkages between SDG 1 and SDG 2. Those linkages can be seen in Figure 3 as well. Sentences about SDG 1 are classified as SDG 2 with the second highest probability. The same is true vice versa for sentences about SDG 2.

Although the labels for SDGs 4, 5, 8 and 10 have the second highest probability after SDG 10, the label for SDG 16 does not work as well as the other labels, as the probability of sentences actually being about SDG 16 is not as high. Additionally, the correlations between SDGs 16, 11 and 14 are not captured by the label. The label for SDG 17 also does not show the highest probability for the sentences that specifically talk about it; instead, it has similar probabilities for all SDGs.

The validation results shown in Figure 3 highlight the importance of using a multi-class model instead of a single-class model. Our findings, which are consistent with the linkages provided by the UN (United Nations, 2016a), demonstrate that sentences often relate to multiple SDGs instead of only one. For example, SDG 17 is about partnership for the goals and is therefore indirectly included in all other goals, which is reflected in our validation results. These results underscore that simply counting the word ‘SDG’ along with the numbers 1 through 17 would not adequately capture the semantics of sentences that discuss multiple SDGs without explicitly mentioning the term ‘SDG’. Our personal experience in reviewing the gold standard labels also supports these findings, as humans would often assign two SDG labels to a single sentence.

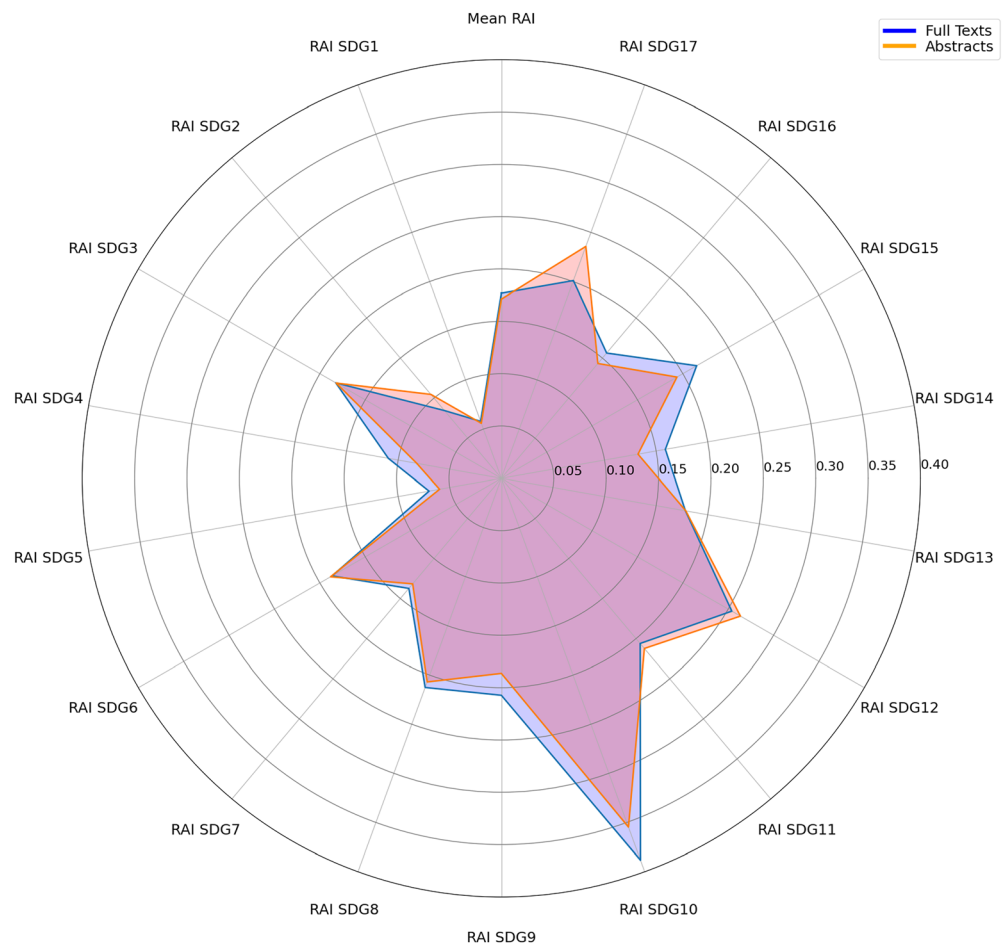
5 | ABSTRACTS VERSUS FULL TEXTS

To strengthen and validate our study's conclusions, we expanded our analysis from abstracts to full texts. This expansion was undertaken to solidify the robustness of our findings, as relying solely on abstracts might not capture the full depth of the research related to SDGs. Our comparison revealed that the analysis of abstracts alone yielded results comparable to those obtained from full texts. For this comparison, we curated a distinct dataset focused exclusively on open-access publications from 2015 to 2022. Based on the 13,138 abstracts included in our initial analysis, we identified 8,503 open-access articles using Web of Science metadata. Of these, we successfully downloaded 7,257 full-text articles. After obtaining the full-text articles, the next step involves preparing these documents for subsequent processing, a crucial stage before employing the zero-shot text classification, as described in the next section.

The ‘From PDFs to raw text data’ section in the Appendix outlines our process of extracting text from research articles using the Visual Layout (VILA) tool. We successfully retrieved text from 7,229 PDFs, converting this data into CSV format for more efficient processing. Through subsequent cleaning steps, which included filtering for documents with a tagged main body and verifying the language as English using the Python Langdetect library, we prepared a final dataset of 7,090 articles for comparative analysis. Similar to the methodology used for the abstracts, we applied zero-shot text classification to the full texts of these articles to derive the RAI. To provide a clear comparison between the two datasets, Figure 4 presents the RAIs for the abstracts, from the 7,090 articles, and their corresponding extracted full texts. This comparative analysis allows us to evaluate any discrepancies or patterns emerging from the abstracts versus the full text, thereby enriching our understanding of the distribution of research attention across the SDGs.

As illustrated in Figure 4, we note that SDG 10 consistently exhibits the highest probability of attention across the abstracts and complete documents. This observation aligns with our validation findings, suggesting that SDG 10 is often referenced, even in contexts primarily discussing other SDGs.

FIGURE 4 Comparison of the research attention index (RAI) for all SDGs, utilizing both abstracts and full articles.



We observed a significant similarity between results derived from abstracts and full texts, indicating that abstracts alone have the capacity to effectively encapsulate the core message of the articles. The average difference between the full texts and abstracts is minor, at just 0.0056 points. The most noticeable deviations were found for SDGs 10 and 17, with differences of 0.0334 and 0.035 points respectively. These findings, as illustrated in Figure 4, have notable implications not only for the precision of analysis, but also for computational efficiency and the potential for significant time savings.

These findings underscore the value of our initial decision to focus on abstracts in our analysis, as our goal is to pinpoint sections where the scientific community explicitly discusses and gives attention to the SDGs. In fact, our results reaffirm the function of abstracts as concise summaries of the most salient points of an article. We observe that the RAI values are generally similar for abstracts compared to full articles, thereby validating our approach.

6 | RESULTS

6.1 | Research attention index

We propose the novel RAI that enables us to compare the relationship between SDGs and research attention at a country level. To achieve

this, we utilize the geographical distribution across countries explored in section ‘Geographical Patterns’ and combine it with the results of our zero-shot classification. We compute the mean probabilities of all sentences in an abstract and title, regarding the SDGs being discussed for a country, and use this as a proxy for the research attention or the SDG discourse in a given country. The zero-shot classification results yield a probability that can be interpreted as a natural index, with a spectrum from zero (indicating no research attention to an SDG in a country) to one (signifying maximum attention to an SDG).

Furthermore, it is important to consider that the RAI we propose here is a relative measure and may not accurately reflect the absolute levels of research attention an SDG receives within a country. Factors such as publication accessibility, language barriers, and regional research priorities can influence the distribution of research articles across countries. Consequently, the index should be perceived as a comparative tool to gauge research activities connected to SDGs across different nations, rather than an absolute measure of research attention an SDG receives.

Figure 5 presents a boxplot to summarize the distribution of research attention across all countries in our dataset with five or more research articles. It is essential to note that we reduced our dataset to 5,646 abstracts, but the results are qualitatively similar for the mean values of research attention provided for the full dataset, as illustrated in Figure A1 in the Appendix. This finding suggests that there is no

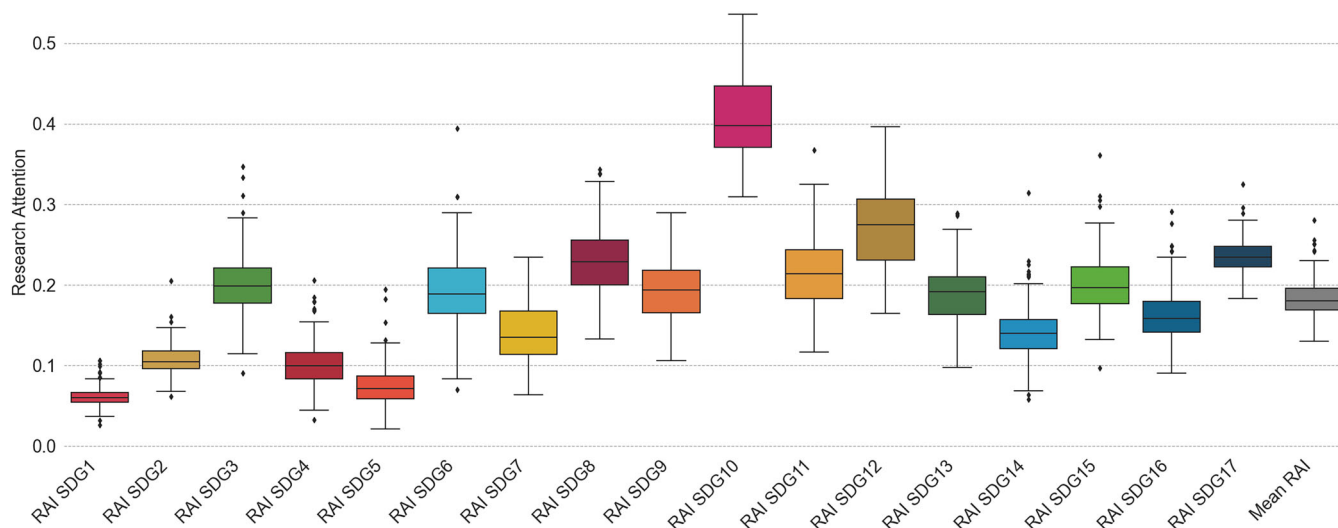


FIGURE 5 Boxplots representing research attention, based on the mean for countries with at least five research articles.

apparent bias in the datasets with respect to the distribution of research attention. When comparing the distribution of mean values over countries obtained from the zero-shot classification (Figure 5), we find that SDGs 1, 2, 4, and 5, which pertain to *no poverty*, *zero hunger*, *quality education*, and *gender equality*, respectively, had the lowest research attention index, with less than 0.11 on average. These goals appear to be less prominent in the scientific discourse of articles indexed in Web of Science.

On the other hand, we find that discourses on SDGs 3, 6, 8, 9, 11, 13, and 15, which include the goals *good health and well-being*, *clean water and sanitation*, *decent work and economic growth*, *industry, innovation, and infrastructure*, *sustainable cities and communities*, *climate action*, and *life below water*, have a similar research attention score on average ranging between 0.18 and 0.22. In contrast, the top two SDGs in terms of research attention are SDG 10—*reduced inequalities* and SDG 12—*responsible consumption and production*. Overall, these findings underscore the reality that research attention towards the SDGs is not uniformly distributed across countries, and there is substantial variation in the level of research focus accorded to different SDGs.

A further comparison of the RAI with the number of articles for each SDG (Table A2 in the Appendix) highlights both similarities and differences in representation of SDGs across research articles. The top five SDGs in the RAI are SDG 10, SDG 12, SDG 17, SDG 11, and SDG 8, while the top five SDGs with the highest number of articles from simple counting are SDG 3, SDG 6, SDG 2, SDG 11 and SDG 7.

Comparing these two methods highlights some similarities and differences in the representation of each SDG in the research landscape. For instance, both methods indicate high representation for SDG 11—*sustainable cities and communities* and medium-to-high representation for SDG 3—*good health and well-being*, while SDG 16—*peace, justice, and strong institutions* consistently ranks low. Contrastingly, SDGs 10—*reduced inequalities* and 12—*responsible consumption and production* rank highly in zero-shot results but lower in article

mentions. SDG 6—*clean water and sanitation* ranks second in article mentions but eighth in zero-shot results. Considering that the zero-shot classification is a more sophisticated method providing deeper insights, it offers a more nuanced understanding of research attention related to each SDG by capturing the broader context and relationships between SDGs, resulting in a more comprehensive understanding of the research landscape. This comparison emphasizes the importance of examining research coverage from multiple perspectives to better understand attention distribution across SDGs, guiding future efforts to address gaps and enhance overall understanding.

6.2 | (Non)-linear relationship between research attention and SDGs

In the next step, we investigate the relationship between scientific discourse and actual development towards the SDGs. There are two plausible scenarios to consider: first, countries facing challenges in achieving the SDGs may be more likely receive attention towards them in the scientific community. Alternatively, it is possible that countries with higher levels of scientific discourse on a particular SDG may be more likely to achieve progress towards that SDG.

In our initial analysis, we delve into the official SDG Index scores and present our findings through Figure A2 in the Appendix, which features a boxplot for each SDG based on the official SDG Index score (United Nations, 2021) for countries within our sample. Several noteworthy observations emerge from these boxplots. Notably, many countries perform well in regards to SDGs 1, 4, 12, and 13, with the majority of countries achieving a median SDG Index score of 80% or higher. However, the remaining SDGs exhibit greater variability in SDG Index scores, indicating that achieving the development goals by 2030 poses a significant challenge for most countries in our sample. For instance, SDGs 5, 7, and 9 through 11 demonstrate the greatest

variance in the average SDG Index score. However, it is important to note that interpretations of the SDG Index score must be approached with caution as they have been subject to criticism (Diaz-Sarachaga et al., 2018; Kroll et al., 2019).

To investigate the degree of attention or focus that the issues related to each SDG receive in a country, we analyze the association between the RAI and their respective average SDG Index scores. Figure 6 illustrates a scatterplot of the overall research attention, categorizing countries based on their income status into four classes: Low-Income Countries (LIC), Lower-Middle-Income.

Countries (LMIC), Upper-Middle-Income Countries (UMIC), and High-Income Countries (HIC). The scatterplot is plotted against the SDG Index Score, which is determined as the average of the official SDG Index scores over the period from 2015 to 2022. Moreover, using the averages between 2015 and 2022 has the advantage that it smoothes out possible year-to-year fluctuations in both the RAI and the SDG Index scores. This mitigates the impact of potential data irregularities or anomalies in specific years. As a result, we can provide a more robust and stable estimate of the relationship between research attention and SDG progression, rather than focusing on a specific point in time. By examining the scatter plot, we aim to identify any patterns or trends in the amount of attention, in terms of research output, that each SDG is receiving and the SDG achievement in a particular country. This analysis can help shed light on whether more attention is associated with higher SDG Index scores or vice versa. Additionally, it can also reveal non-linear relationships or varying degrees of correlation between research output and different SDGs Index scores, which can be valuable information for guiding future research and policy-making efforts in the field of sustainable development.

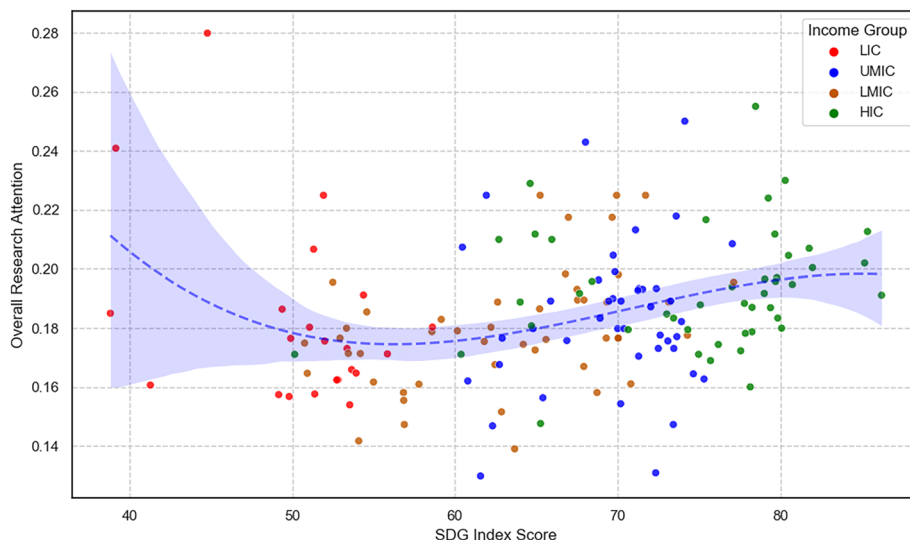
To conduct this analysis, we determine the optimal polynomial order for regression using leave-one-out cross-validation. This method helps us identifying the best polynomial order for each pair of SDG research attention and average SDG Index score, ensuring the best fit for the data while minimizing overfitting. In the case of Figure 6, the

blue dashed regression line has been fitted with a polynomial of third order, while the light shaded area represents the 95% confidence band. By integrating these approaches, we can effectively analyze the relationship between research attention and SDG achievement, offering insights that can guide both research and policy directions in sustainable development. The results become particularly intriguing when considering the attention related to SDGs, their respective average scores, and the income group of countries. At an aggregated level, we observe a clear indication of a non-linear relationship, which is generally upward sloping. Research attention appears to exhibit a downward slope for LICs, while it demonstrates a slightly upward trend for LMICs, that increases for UMICs and HICs.

This observation underscores the necessity for further investigation into the factors contributing to these varying trends across income groups and their influence on progress towards achieving the SDGs. To obtain a more in-depth understanding, we will examine the disaggregated level, focusing on individual SDGs, from SDG 1 to SDG 17, as depicted in Figure 7. The optimal order of the fitted regression polynomial is observed to fluctuate across the SDGs. While a linear relationship (order 1) is evident for SDGs 1, 2, 4, 6, 7, 9, 14, and 17, multiple instances of non-linear relationships are also present. SDGs 5, 8, 11, 13, 15, and 16 appear to follow a quadratic relationship, whereas SDGs 3, 10, and 12 are described by a higher-order polynomial.

To examine the relationship between the RAI and the SDGs in greater detail, we employed Kendall's rank correlation coefficient as our analysis method, as illustrated in Table 2. This table showcases the Kendall rank correlation coefficients between research attention and the official average SDG Index score per country (United Nations, 2021). The table is organized into five columns, representing the overall correlation and correlations across four distinct income levels: LIC, LMIC, UMIC, and HIC. We chose Kendall's Tau over Pearson's correlation coefficient, as the latter presumes a normal distribution of the variables and a linear relationship—assumptions that do not align with our data. Moreover, we favored Kendall's Tau over

FIGURE 6 Scatter plot illustrating the relationship between mean RAI and SDG Index score. Countries are categorized based on their income status into four classes: Low-income countries (LIC), lower-middle-income countries (LMIC), upper-middle-income countries (UMIC), and high-income countries (HIC). The blue dashed line represents the optimal order of the fitted regression polynomial using a leave-one-out cross-validation (order 3), while the light shaded area represents the 95% confidence band.



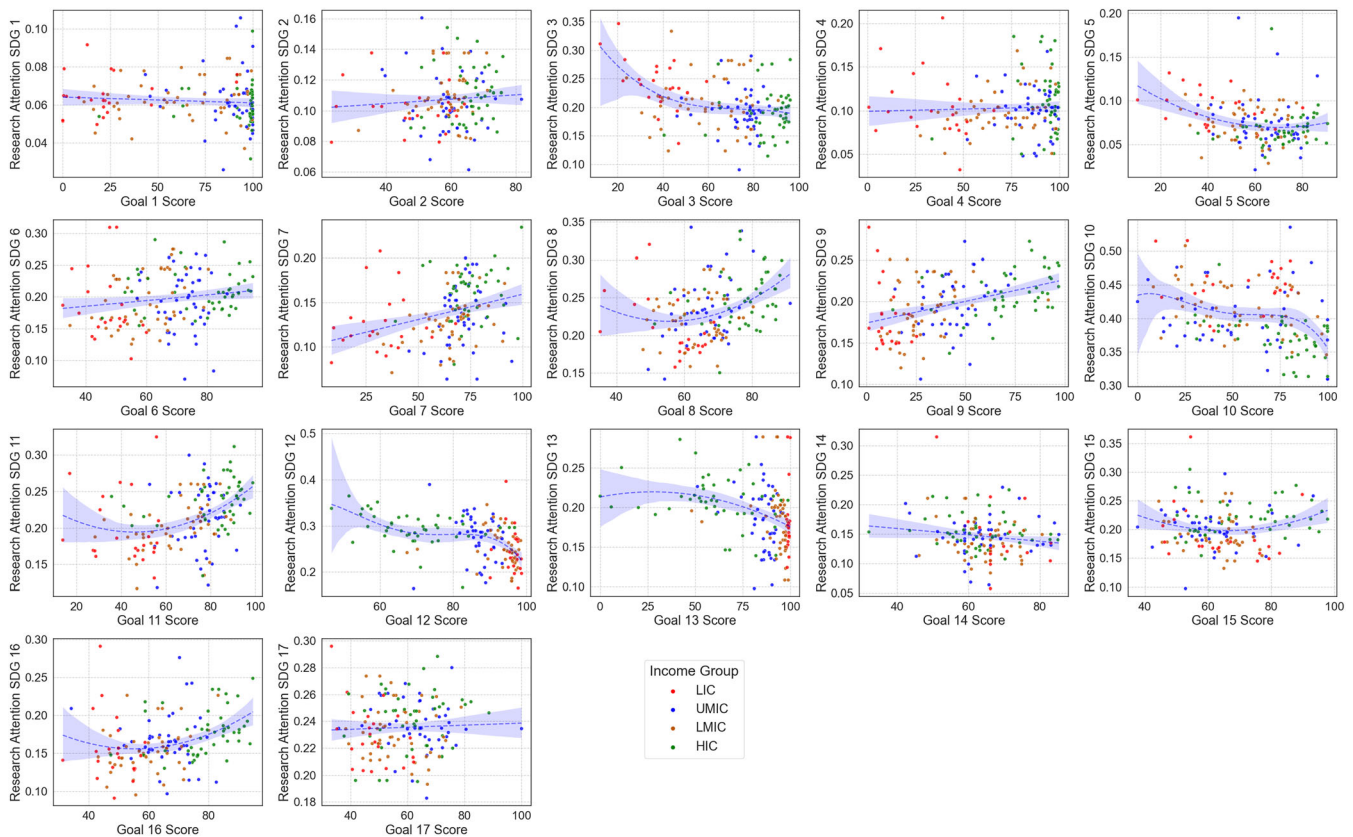


FIGURE 7 Scatter plots illustrating the relationship between research attention and individual SDG index scores. Countries are categorized based on their income status into four classes: Low-income countries (LIC), lower-middle-income countries (LMIC), upper-middle-income countries (UMIC), and high-income countries (HIC). The blue dashed line represents the optimal order of the fitted regression polynomial using a leave-one-out cross-validation, while the light shaded area represents the 95% confidence band.

Research attention versus SDG Index score	Overall	LIC	LMIC	UMIC	HIC
SDG Index Score	0.23***	-0.11	0.28***	0.03	0.22**
SDG 1	-0.13**	0.13	0.06	-0.19*	0.05
SDG 2	0.09*	0.02	0.11	-0.06	0.11
SDG 3	-0.19***	-0.34**	-0.05	-0.02	0.1
SDG 4	0.05	-0.24	0.04	-0.05	-0.06
SDG 5	-0.14***	-0.35**	-0.05	-0.15	0.25**
SDG 6	0.13**	-0.23	0.21**	0.03	0.14
SDG 7	0.25***	0.08	0.29***	0.12	0.12
SDG 8	0.23***	-0.26	0.04	0.22**	0.19*
SDG 9	0.29***	-0.26*	0.23**	0.18*	0.32***
SDG 10	-0.25***	0.11	-0.26***	-0.06	-0.32***
SDG 11	0.28***	-0.05	0.23**	0.09	0.22**
SDG 12	-0.35***	-0.17	-0.27***	-0.22**	-0.24**
SDG 13	-0.27***	0.13	-0.3***	-0.16	-0.2*
SDG 14	-0.08	-0.2	-0.06	0.05	-0.12
SDG 15	0.01	-0.28*	-0.11	-0.09	0.12
SDG 16	0.24***	-0.12	0.23**	-0.02	0.27***
SDG 17	0.04	-0.23	-0.0	-0.12	0.1

TABLE 2 Kendall rank correlation between RAI and SDG index score per country.

* $p < .1$; ** $p < .05$; *** $p < .01$.

Spearman's correlation coefficient, as Kendall's Tau is generally seen as more robust, especially when handling smaller sample sizes or datasets with a significant number of tied ranks.

Table 2 shows varying degrees of correlation between RAI and the different SDGs overall and across different income classifications. The overall correlation between research attention and the SDG Index Score is positive and statistically significant, indicating that, in general, higher research attention is associated with a higher SDG Index score, and vice versa. A fact, that was also visible as a general indication on Figure 6. However, when looking at the correlations for each individual SDG, we observe mixed results. On the overall level, we find a significant positive correlation for SDGs 2, 6, 7, 8, 9, 11 and 16, which is in line with the overall trend. However, we find a significant negative correlation for SDGs 1, 3, 5, 10, 12, 13, 14. For SDGs 4, 15, and 17, the RAI does not exhibit a significant correlation with the SDG Index scores.

The overall correlation between RAI and SDG Index Score is significantly positive (0.23**). However, for SDG 1—*no poverty*, we observe a contrary effect (−0.13**). This could imply that even in areas where research attention is high, poverty might remain a significant issue. On the other hand, it could mean that the awareness for the problems is high and more research is being conducted about these countries. This could stem from the specific focus of the research being conducted, or perhaps the benefits of heightened research attention have not yet sufficiently mitigated poverty levels. Furthermore, in order to alleviate the dire situations caused by high poverty levels, it is imperative to dedicate more research attention to these countries. This focus on research could help improve the conditions and ultimately reveal a negative correlation. Additionally, the negative correlation with SDG 3—*good health and well-being* (−0.19***), indicates that an increase in research attention could be associated with factors such as elevated stress levels, inadequate work-life balance, or other health-related concerns.

As suggested by Figure 2, SDGs 7—*affordable and clean energy*, and SDG 13—*climate action*, frequently co-occur in academic literature, underlining their importance for an effective transition towards achieving the 1.5-degree goal. However, intriguingly, we observe contrasting correlations for these two SDGs. SDG 7 exhibits a strong positive correlation (0.25***) between research attention and SDG achievement, suggesting that increased research focus may directly contribute to advancements in clean energy technology or strategies, thereby driving progress in this SDG. The progression of clean energy relies heavily on strategic investments in this sector, fostering increased attention in international discourse and consequently attracting greater research focus. While it is evident that research activities can stimulate advancements in SDG 7, it is equally plausible that progress made in SDG 7 can attract further research attention, demonstrating the bidirectional and dynamic nature of this relationship. In contrast, SDG 13 displays a strong negative correlation with research attention (−0.27***). However, it is worth noting that wealthier countries often score lower overall on SDG 13, despite possessing the resources and capabilities necessary to combat the climate crisis. This highlights the critical need to understand the relationships

between income levels, research attention, and progress towards not just climate action but all other SDGs as well. When examining the disaggregated effects, we can observe some interesting patterns. Particularly, in the LIC group, we generally observe either negative correlations or statistically insignificant correlations between research attention and SDG Index scores. For instance, the relationship between research attention and SDG 3—*good health and well-being* and SDG 5—*good health and well-being* is notably negative (−0.34** and −0.35**, respectively). These correlations suggest that in these countries, an increase in research attention does not necessarily lead to improvements in health or gender equality and could potentially even widen existing disparities. On the other hand, increasing disparities in health or gender equality could also lead to heightened research attention, which could be an attempt to understand the root cause of these disparities.

For LMICs, the correlations vary considerably across SDGs. Notably, there are strong positive correlations for SDG 7—*affordable and clean energy* (0.29***) and SDG 11—*sustainable cities and communities* (0.23**), suggesting that research attention in these countries might be particularly beneficial for these areas. Additionally, advancements in these areas, as indicated by the SDG Index Scores, could elevate their prominence and perceived significance within the research community. This heightened awareness could, in turn, stimulate increased research attention towards these SDGs. Conversely, significant negative correlations exist with SDG 10—*reduced inequalities* (−0.26***) and SDG 13—*climate action* (−0.30***). This implies that growing inequalities and the escalating urgency of climate change might be attracting more research focus, as scholars strive to understand and address these pressing issues.

For UMICs, the results are generally weaker, with fewer statistically significant findings. However, SDG 8—*recent work and economic growth* and SDG 9—*industry, innovation, and infrastructure* (0.22** and 0.18*, respectively) show a positive correlation, suggesting that increased research attention might have a positive impact on economic growth and innovation in these countries, while economic growth could also provide an environment that provides more funding opportunities.

Finally, HICs predominantly exhibit positive correlations between research attention and SDGs, with the strongest being SDG 9—*industry, innovation, and infrastructure* (0.32***). However, SDG 10—*reduced inequalities* and SDG 12—*responsible consumption and production* (−0.32*** and −0.24**, respectively) have significant negative correlations, suggesting that higher research attention might coincide with increased inequalities and less responsible consumption and production patterns.

Interestingly, we observe a trend towards a non-linear relationship between research attention and SDG Index scores, particularly for SDGs 5 and 9, and to some extent for SDG 8.

For SDG 5—*gender equality*, we find a negative and statistically significant correlation for LICs at −0.35**. This might imply that gender inequality issues could be attracting increased research attention as scholars seek to understand and propose solutions for these pressing issues. However, heightened research attention does not

necessarily lead to improved gender equality. One explanation could be entrenched gender disparities or restricted access to education and opportunities for women. Conversely, the correlation for HICs is positive and significant (0.25***), which implies that increased research attention is associated with better gender equality outcomes. The disparity between income groups may stem from variations in socio-economic and cultural contexts, which could influence the impact of research attention on gender equality.

For SDG 9—*industry, innovation, and infrastructure*, we observe a reversed pattern. While the overall correlation between research attention and SDG 9 is positive and significant (0.29***), it becomes negative and significant for LICs (−0.26*). This suggests that in low-income countries, increased research attention may not directly translate into advancements in industry, innovation, or infrastructure. This could be attributed to resource constraints, underinvestment in these sectors, or hurdles in implementing research outcomes. Moreover, the association between increased research attention and a lower score appears to be more pronounced in these particular countries. However, in LMICs, UMICs, and HICs, the correlations are positive and significant (0.23**, 0.18*, and 0.32***, respectively). This implies that in these countries, heightened research attention tends to be associated with improved performance in achieving SDG 9. Particularly in HICs, it seems reasonable to assume that they have the necessary resources and infrastructure to leverage research findings effectively, thereby driving advancements in industry, innovation, and infrastructure and their successes become subjects of interest for researchers aiming to understand and replicate these advancements.

Lastly, we note a shift in sign for SDG 8—*decent work and economic growth*, moving from a negative correlation for LICs (−0.26), albeit non-significant, to a significant positive correlation for UMICs and HICs (0.18* and 0.32***, respectively). This negative correlation in LICs could indicate that heightened research attention does not necessarily result in job creation or economic growth. This could be due to a misalignment between the focus of research and the needs of the local economy, or difficulties in implementing research findings. Conversely, for UMICs and HICs, the positive correlation might suggest that increased research attention tends to foster innovation, create jobs, and stimulate economic growth. These countries, with their more developed economies, robust institutions, and superior resources and infrastructure, are often better positioned to actualize the benefits of research.

In summary, our findings indicate substantial variation in research attention across different income groups and SDGs. This variation could reflect differing capacities to utilize research findings, differing research priorities, and unequal distribution of research resources. It could also reflect broader socio-economic trends and disparities. For instance, wealthier countries with more resources might be better equipped to translate research attention into progress towards SDGs. Their success in these areas might, in turn, draw further research attention. In contrast, in lower-income countries, other factors such as poverty, lack of infrastructure, and political instability might hinder the positive impacts of research. Paradoxically, these challenges could also attract increased research attention, as there is an urgent need to address these pressing issues.

7 | DISCUSSION

Overall, two contrasting theories could potentially explain the relationship between research attention towards SDGs and the actual SDG Index scores. The first theory suggests a negative relationship, suggesting that SDGs with lower scores receive more attention. The second theory suggests a positive correlation, implying that countries with higher SDG Index scores may draw more research and knowledge sharing attention. We observe evidence supporting both theories in our data, depending on the specific SDG. In the following, we delve into broader explanations for our results, building on the more detailed discussions in section ‘(Non)-Linear Relationship between Research Attention and SDGs’.

Some explanations for the first theory can be derived from theories about media coverage and human reactions to it. Soroka et al. (2019) suggest that humans tend to react more strongly to negative news coverage, a phenomenon labeled as ‘negative biases’ (Rozin & Royzman, 2001). Similarly, low SDG Index scores might receive more research coverage, while those with high SDG Index scores may not be as prominently featured, resulting in a negative relationship between research attention towards SDGs and the SDG Index scores.

Furthermore, general psychological concepts suggest that people tend to be more sensitive to negative information (Baumeister et al., 2001). Thus, low-scoring SDGs might trigger stronger emotional responses and consequently attract more attention. Researchers might be more moved by lower SDG Index scores, thereby concentrating their research efforts on countries and SDGs with such scores.

Countries with lower SDG Index scores may indeed attract more funding and resources due to the perceived potential for improvement. This increase in financial support can lead to more attention being directed towards these countries and the specific challenges indicated by their lower scores. Recognition of the need for improvement can stimulate interest and investment from a variety of stakeholders, including governments, development agencies, NGOs, and philanthropies. As a result, research, initiatives, and collaborations focused on addressing the challenges and promoting sustainable development in these countries can receive heightened attention and support (OECD, 2002).

While we do not find such a negative relationship for the overall correlation between SDG Index Score and research attention, the arguments made above are plausible for some SDGs. In our analysis we find support for those theories with negative correlations between the SDG Index Scores and research attention for SDGs 1, 3, 5, 10, 12, 13, and 14. Thus, for these SDGs, countries with lower scores might attract more funding and attention, or the psychological explanation or any of the reasons stated above might hold true.

On the other hand, there is an alternative scenario suggesting a positive correlation between research attention and the SDG Index score. This theory implies that countries with higher SDG Index scores may receive greater attention, particularly in terms of research and knowledge sharing. Several explanations support this viewpoint:

Firstly, countries with higher SDG Index scores are generally more developed and possess greater resources, including technological,

financial, and human capital (United Nations IATF, 2022). As a result, these countries are better equipped to invest in sustainable development, dedicating significant resources to research and development. Their commitment to sustainable development and their ability to collaborate with international partners contribute to an increased focus on the topic (King, 2004).

By leveraging their resources and engaging in research and development activities, countries with higher SDG Index scores demonstrate their commitment to advancing SDGs. This commitment and investment draw attention from researchers, policymakers, and practitioners who seek to understand and learn from their successes. As a result, these countries become focal points for research and knowledge sharing, providing valuable insights and best practices that can be replicated in other contexts (OECD, 2021).

Secondly, countries with higher SDG Index scores serve as compelling examples of success in implementing sustainable development policies and practices. Researchers, policymakers, and practitioners are often drawn to studying these countries to gain insights into the factors that contributed to their achievements, with the aim of replicating their success in other contexts (Sachs et al., 2019; Sachs et al., 2022; Sachs, Kroll, et al., 2021; Sachs, Schmidt-Traub, et al., 2021).

Another driving factor for the enhanced attention towards high-scoring countries is the recognition and support they receive from the international community, including development agencies, non-governmental organizations, and academic institutions. These countries often serve as role models in various domains, such as digital transformation, thereby setting benchmarks for other nations. This recognition and admiration lead to enhanced research collaborations, increased funding opportunities, and greater knowledge sharing, all focused on promoting sustainable development practices (OECD, 2021).

Furthermore, countries with higher SDG Index scores often wield more influence in steering global sustainable development agendas and priorities. Their successful experiences and perspectives command greater attention in international forums and policy discussions, as they offer valuable insights and guidance for other nations (OECD, 2021).

Our results indicate a positive overall correlation between research attention and the SDG Index scores, which lends more weight to the arguments posited above. At the individual SDG level, a positive correlation can be found for SDGs 2, 6, 7, 8, 9, 11 and 16 as well. Consequently, for these SDGs, reasons stated above, such as being a role model as a high-scoring country or attracting more attention in international forums or policy discussions, might hold true.

Initially, we assumed that the reasons for a negative relationship would be more credible. However, our results reveal evidence supporting both sides of the debate, preventing us from unequivocally favoring one over the other.

8 | CONCLUSION

In conclusion, this study has provided a new avenue for the exploration of the global research landscape surrounding SDGs. In addressing

our first research question, we focused on the effective utilization of advanced NLP techniques, particularly pre-trained language models and zero-shot text classification, for analyzing large-scale scientific textual data related to the SDGs. We demonstrated that these techniques allow for a nuanced and efficient analysis, uncovering insights that are often hidden in traditional analysis methods. This approach has enabled us to better understand the complex dynamics within SDG-related research, providing a valuable tool for future studies in this field.

In response to our second research question regarding the novel insights these advanced NLP techniques could reveal, our study identifies a significant gap in the current literature: the underutilization of NLP methods in large-scale analysis of SDG-related scientific discourse. Despite a growing body of research using NLP, there is a need for more advanced textual analysis that is capable of capturing the SDG discourse within the scientific community. Using the power of zero-shot text classification as a modern NLP tool and the introduction of the RAI, we have provided a fresh perspective on the dynamics of SDG-focused scientific discourse, its relationship with SDG achievement and the correlation across different income levels in countries. This offers valuable insights for researchers, policymakers, and stakeholders involved in sustainable development efforts. In relation to the third research question, our use of the RAI and NLP analysis aims to uncover the connections between SDG index scores and research attention for each specific SDG. This valuable information can be utilized by policymakers to prioritize their attention towards particular SDGs and adjust their strategies accordingly. By identifying interesting relationships within scientific literature, policymakers can effectively concentrate their efforts on relevant studies, especially when it comes to unexpected and non-linear relationships like those observed for SDGs 5, 8, and 9. This highlights the potential of scientific literature to offer insights that may not be readily apparent, ultimately helping in improving policy implications. In summary, our study's primary contribution resides in equipping policymakers with knowledge and presenting them with a tool to efficiently navigate the vast amounts of scientific literature.

Nevertheless, it is crucial to acknowledge the limitations of our analysis and to use the RAI as a tool for relative comparison rather than an absolute measure of research attention. Future research should concentrate on addressing the limitations identified in this study. One such limitation is the complexity in comparing the composite measurement of the SDG Index scores and the RAI. Our analysis offers a composite measure of research attention spanning the publication years of 2015–2022, paralleling the UNs' SDG Index scores, which are computed by combining indices from the same time period. Consequently, our comparison reflects an aggregation of data from several years, as opposed to a time-series analysis. Although this method provides a general trend, a more granular, year-by-year comparison could yield a more in-depth analysis and provide a better understanding of the dynamic involved. At present, we have been able to identify a correlation between research attention and SDG achievement. However, the directionality of this relationship remains unclear. Therefore, a more detailed analysis could shed light on this complex

interaction and enhance our understanding of its intricate mechanisms. However, obtaining the required data is challenging due to the complex structure of the SDG indices and the current non-stationary behavior of much of the data makes this level of analysis unattainable at present.

In addition, our data source was largely based on the abstracts of metadata provided by the Web of Science. While Web of Science is a comprehensive and well-regarded database for scientific research, it does not capture the entire global research output related to SDGs. There are potential other repositories such as ArXiv, various policy briefs, pre-print articles, and other machine-readable text formats could significantly contribute to the SDG research landscape. The used method of zero-shot text classification is robust and can be applied to any machine-readable text format. However, our analysis did not take these resources into account. Thus, future research could build on our approach, extending it to broader range of sources. This could lead to the creation of a more generalized and comprehensive version of the RAI, contributing a more holistic perspective on the global research attention devoted to the SDGs.

Moreover, while we opted for UN Indicator scores to bring transparency and simplicity to our analysis, alternative measures could also provide insightful and relevant perspectives on SDG achievement. Future research could consider incorporating other indicators, which provide a different perspective on SDG achievement.

Lastly, our analysis did not account for several potential confounding factors that could influence the relationship between research attention and SDG performance. Factors such as government efficiency, research funding, and socio-economic context may all significantly impact this relationship. Thus, future research should seek to better understand the dynamics between research attention and SDG achievement by investigating the underlying mechanisms and factors driving the observed variations in correlations across different income groups. By deepening our understanding of these complex dynamics, we can more effectively develop and implement policies that facilitate progress towards achieving the SDGs across all income levels. We hope that our findings will stimulate a wider discourse on the correlation between research attention and SDG performance, as well as the factors driving this relationship.

AUTHOR CONTRIBUTIONS

Christoph Funk: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Elena Tönjes:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Ramona Teuber:** Conceptualization, Supervision and Writing – Review & Editing. **Lutz Breuer:** Conceptualization, Funding Acquisition, Project Administration Supervision, Writing – Review & Editing.

ACKNOWLEDGMENTS

Christoph Funk, Ramona Teuber and Lutz Breuer acknowledge funding by the German Academic Exchange Service (DAAD) from funds of the Federal Ministry for Economic Cooperation (BMZ), SDGNexus

Network (Grant No. 57526248), program “exceed—Hochschulexzellenz in der Entwicklungszusammenarbeit”. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

ORCID

Christoph Funk  <https://orcid.org/0000-0001-7388-4415>

REFERENCES

- Angin, M., Taşdemir, B., Yılmaz, C. A., Demiralp, G., Atay, M., Angin, P., & Dikmener, G. (2022). A roberta approach for automated processing of sustainability reports. *Sustainability*, 14(23), 16139.
- Bali Swain, R., & Ranganathan, S. (2021). Modeling interlinkages between sustainable development goals using network analysis. *World Development*, 138, 105136.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370.
- Bellantuono, L., Monaco, A., Amoroso, N., Aquaro, V., Lombardi, A., Tangaro, S., & Bellotti, R. (2022). Sustainable development goals: Conceptualization, communication and achievement synergies in a complex network framework. *Applied Network Science*, 7(1), 14.
- Beltagy, I., Cohan, A., & Lo, K. (2019). Scibert: Pretrained contextualized embeddings for scientific text. *CoRR*, <https://arxiv.org/abs/1903.10676>. Last accessed Feb. 28, 2022.
- Bennich, T., Weitz, N., & Carlsen, H. (2020). Deciphering the scientific literature on sdg interactions: A review and reading guide. *The Science of the Total Environment*, 728, 138405.
- Chang, I.-C., Yu, T.-K., Chang, Y.-J., & Yu, T.-Y. (2021). Applying text mining, clustering analysis, and latent dirichlet allocation techniques for topic classification of environmental education journals. *Sustainability*, 13(19), 10856.
- Davison, J. (2020). Zero-shot learning in modern nlp. Last accessed Feb. 28, 2022.
- Decouttere, C., De Boeck, K., & Vandaele, N. (2021). Advancing sustainable development goals through immunization: A literature review. *Globalization and Health*, 17(1), 95.
- Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *CoRR*, <https://arxiv.org/abs/1810.04805>. Last accessed Feb. 28, 2022.
- Diaz-Sarachaga, J. M., Jato-Espino, D., & Castro-Fresno, D. (2018). Is the sustainable development goals (sdg) index an adequate framework to measure the progress of the 2030 agenda? *Sustainable Development*, 26(6), 663–671.
- Fuso Nerini, F., Sovacool, B., Hughes, N., Cozzi, L., Cosgrave, E., Howells, M., Tavoni, M., Tomei, J., Zerriffi, H., & Milligan, B. (2019). Connecting climate action with other sustainable development goals. *Nature Sustainability*, 2(8), 674–680.
- Hackl, A. (2018). Mobility equity in a globalized world: Reducing inequalities in the sustainable development agenda. *World Development*, 112, 150–162.
- King, D. A. (2004). The scientific impact of nations. *Nature*, 430(6997), 311–316.
- Kroll, C., Warchold, A., & Pradhan, P. (2019). Sustainable development goals (sdgs): Are we successful in turning trade-offs into synergies? *Palgrave Communications*, 5(1), 140. <https://www.nature.com/articles/s41599-019-0335-5>
- Le Blanc, D. (2015). Towards integration at last? The sustainable development goals as a network of targets. *Sustainable Development*, 23(3), 176–187.

- Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., & Zettlemoyer, L. (2019). BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *CoRR*, <https://arxiv.org/abs/1910.13461>. Last accessed Feb. 28, 2022.
- OECD. (2002). *Foreign direct Investment for Development*.
- OECD. (2021). *Development Co-operation report 2021*.
- Pedercini, M., Arquitt, S., Collste, D., & Herren, H. (2019). Harvesting synergy from sustainable development goal interactions. *Proceedings of the National Academy of Sciences of the United States of America*, 116(46), 23021–23028.
- Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, 5(4), 296–320.
- Sachs, J., Kroll, C., Lafortune, G., Fuller, G., & Woelm, F. (2021). *Sustainable development report 2021*. Cambridge University Press.
- Sachs, J., Kroll, C., Lafortune, G., Fuller, G., & Woelm, F. (2022). *Sustainable development report 2022*. Cambridge University Press.
- Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., & Fuller, G. (2019). This work is licensed under a creative commons attribution 4.0 international license. In *Sustainable Development report 2019*. Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).
- Sachs, J., Schmidt-Traub, G., Kroll, C., Lafortune, G., Fuller, G., & Woelm, F. (2021). *Sustainable development report 2020: The sustainable development goals and Covid-19 includes the SDG index and dashboards*. Cambridge University Press.
- Sebestyén, V., Domokos, E., & Abonyi, J. (2020). Focal points for sustainable development strategies-text mining-based comparative analysis of voluntary national reviews. *Journal of Environmental Management*, 263, 110414.
- Shen, Z., Lo, K., Wang, L. L., Kuehl, B., Weld, D. S., & Downey, D. (2021). Vila: Improving structured content extraction from scientific pdfs using visual layout groups. Last accessed Feb. 28, 2022.
- Smith, T. B., Vacca, R., Mantegazza, L., & Capua, I. (2021). Natural language processing and network analysis provide novel insights on policy and scientific discourse around sustainable development goals. *Scientific Reports*, 11(1), 22427.
- Socher, R., Ganjoo, M., Manning, C. D., & Ng, A. (2013). Zero-shot learning through cross-modal transfer. *Advances in Neural Information Processing Systems*, 26. <https://proceedings.neurips.cc/paper/2013/hash/2d6cc4b2d139a53512fb8cbb3086ae2e-Abstract.html>. Last accessed Feb. 28, 2022.
- Soroka, S., Fournier, P., & Nir, L. (2019). Cross-national evidence of a negativity bias in psychophysiological reactions to news. *Proceedings of the National Academy of Sciences*, 116(38), 18888–18892.
- United Nations. (2016a). A nexus approach for the sdgs. Last accessed Feb. 28, 2022.
- United Nations. (2016b). Transforming our world: The 2030 agenda for sustainable development. Last accessed Feb. 28, 2022.
- United Nations. (2021). Sustainable development report 2021 supplementary online materials. Last accessed Feb. 28, 2022.
- United Nations Department of Economic and Social Affairs. (2019). Handbook for the preparation of voluntary national reviews, the 2019 edition. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2016). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2017). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2018). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2019). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2020). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations Economic and Social Council. (2021). Progress towards the sustainable development goals: Report of the secretary-general. Last accessed Feb. 28, 2022.
- United Nations IATF. (2022). Financing for Sustainable Development Report 2022.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. L., Gugger, S., ... Rush, A. M. (2020). Transformers: State-of-the art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: System demonstrations* (pp. 38–45). Association for Computational Linguistics Last accessed Feb. 28, 2022.
- Xu, Y., Li, M., Cui, L., Huang, S., Wei, F., & Zhou, M. (2020). Layoutlm: Pre-training of text and layout for document image understanding. In *Proceedings of the 26th ACM SIGKDD international conference on Knowledge Discovery & Data Mining* (Vol. 1, pp. 1192–1200). Association for Computing Machinery.
- Yin, W., Hay, J., & Roth, D. (2019). Benchmarking zero-shot text classification: Datasets, evaluation and entailment approach. Last accessed Feb. 28, 2022.

How to cite this article: Funk, C., Tönjes, E., Teuber, R., & Breuer, L. (2024). Reading between the lines: The intersection of research attention and sustainable development goals. *Sustainable Development*, 32(5), 4545–4566. <https://doi.org/10.1002/sd.2906>

APPENDIX A

From PDFs to raw text data.

In the world of scientific research, most articles are only available in the PDF format, and extracting text from these files is a crucial first step in any downstream NLP task. However, unlike common formats like txt or csv, accessing information from PDF files is not as straightforward. Many popular Python modules, such as 'PDFminer,' are not layout sensitive and can result in subpar text extraction quality, missing important information from the layout of PDF files. Our experience has shown that applying such non-layout sensitive models to our scientific PDFs can result in entire sentences being parsed incorrectly or being completely missed. To address this issue, several studies have attempted to improve text and layout extraction methods (e.g., Xu et al. (2020)). Utilizing layout information can help distinguish between actual text information in an article and text extracted from tables or figures, which we do not consider relevant for our analysis (Shen et al., 2021).

To extract text and layout information from scientific articles, we opted for a VILA-based model, specifically tailored to scientific texts,

as proposed by Shen et al. (2021). This model fine-tunes a pre-trained language model, such as BERT, and does not require the costly pre-training process. The VILA model is based on the premise that scientific text can be grouped into blocks or lines, which can be identified using layout detection models or rule-based PDF parsing (Shen et al., 2021). The authors Shen et al. (2021) present two different methods for incorporating the group structure, I-VILA (Injecting Visual Layout Indicators) and H-VILA (Visual Layout-guided Hierarchical Model). However, for our text and layout extraction, we opted for the H-VILA block trained on the gtoap2 training set using the layoutLM model (Xu et al., 2020). We chose this approach based on personal assessment of randomly selected articles, which showed slightly better results compared to other combinations.

The model provides a token classification alongside the parsed text, with classifications such as abstract, author, dates, body content, figure, keywords, title, and more. However, we only utilize the body content text blocks for our analysis. While parsing the PDF text yielded satisfactory results, manual cleaning was necessary to improve sentence semantics and text cleanliness for our topic model. Table A1 outlines the cleaning steps we performed.

TABLE A1 Cleaning steps of parsed raw texts.

Problem	Fix
Extra whitespaces	Replace extra whitespaces by single whitespace
Words separated by hyphen	Remove hyphen
Words separated by whitespace	Remove whitespace
Model parsed some ff, fi and if characters as one special character	Replace special double characters by normal characters
Whitespaces between word in sentence and punctuation	Remove whitespace
URLs in text which do not have any semantic meaning	Remove all urls from text
Parsed some sentences character by character with whitespace between them, that is, S E N T E N C E instead of Sentence. Problem mostly occurred with figure subtexts	As most figure subtexts do not have important semantic meaning, we removed such single characters
Very short sentences or single tokens	Filter out sentences with less than 6 words or 50 characters
Coding errors	Filters out sentences with a high percentage of unrecognized words using a spell checker

TABLE A2 Counted SDGs for abstracts of countries with at least five research articles.

SDG	Total count	Number of articles	Average count per article
1. No poverty	127	94	1.35
2. Zero hunger	155	126	1.23
3. Good health and well-being	268	187	1.43
4. Quality education	138	98	1.41
5. Gender equality	81	70	1.16
6. Clean water and sanitation	231	145	1.59
7. Affordable and clean energy	155	118	1.31
8. Decent work and economic growth	93	78	1.19
9. Industry, Innovation and infrastructure	61	50	1.22
10. Reduced inequalities	69	55	1.25

TABLE A2 (Continued)

SDG	Total count	Number of articles	Average count per article
11. Sustainable cities and communities	178	119	1.50
12. Responsible consumption and production	91	82	1.11
13. Climate action	142	117	1.21
14. Life below water	79	44	1.80
15. Life on land	115	81	1.42
16. Peace, Justice and strong institutions	52	33	1.58
17. Partnership for the goals	63	51	1.24
SDG/SDGs	13,574	5646	2.58

Note: Table presents a comprehensive overview of the total number of SDGs counted, the number of abstracts in which each specific SDG appears at least once, and the average number of mentions per abstract.

TABLE A3 Counts by country code.

ISO2	Count	ISO2	Count	ISO2	Count	ISO2	Count	ISO2	Count
CN	825	IN	623	US	309	ZA	265	BR	238
NG	235	GH	203	BD	199	ES	175	ID	169
ET	166	KE	161	GB	160	PK	157	AU	145
JP	122	MY	113	DE	111	IT	103	TZ	103
IR	100	NP	99	RU	94	MX	91	UG	90
CO	86	CA	86	VN	85	UA	81	TH	78
MW	77	PL	71	SE	62	KR	59	RW	59
PH	59	TR	59	TW	58	MZ	56	EG	53
ZW	53	ZM	48	PT	47	PE	47	FR	46
FI	44	CL	44	CM	42	NL	42	SA	40
LK	40	MM	39	KH	39	IE	39	RO	37
NZ	37	EC	36	CD	35	CH	35	AE	34
AF	32	MG	31	GR	31	DK	31	SN	31
NO	30	AT	30	JO	28	SL	28	KZ	26
AR	26	LT	25	CZ	24	HU	23	NA	22
BF	22	SG	22	MN	22	EE	21	HK	20
LA	20	NE	19	MA	19	ML	18	QA	18
IS	18	RS	17	SI	17	LV	17	SK	17
BO	16	DZ	16	FJ	15	GT	15	AO	14
CR	14	SD	14	LB	14	CG	13	BG	13
BT	12	BE	12	BJ	12	BH	12	LR	12
HT	12	HR	11	PS	11	HN	11	KW	11
PG	11	TD	11	UY	11	CF	11	OM	10
CU	10	YE	10	BW	10	GN	9	TN	9
VU	9	JM	9	KG	9	TL	9	PY	8
IL	8	LS	8	LU	8	MU	8	SS	8
CY	7	SO	7	GM	7	LY	6	MO	6
TG	6	TJ	6	NI	6	SZ	6	IQ	6
BI	6	SB	5	MV	5	CI	5	MT	5
AZ	5	DO	5	GE	5	BZ	5	AL	5
TT	5	GY	5	ME	5	AM	5	BY	5

Note: Table shows the ISO2 country code and the corresponding number of research articles included for the Research Attention Index (RAI) in each country.

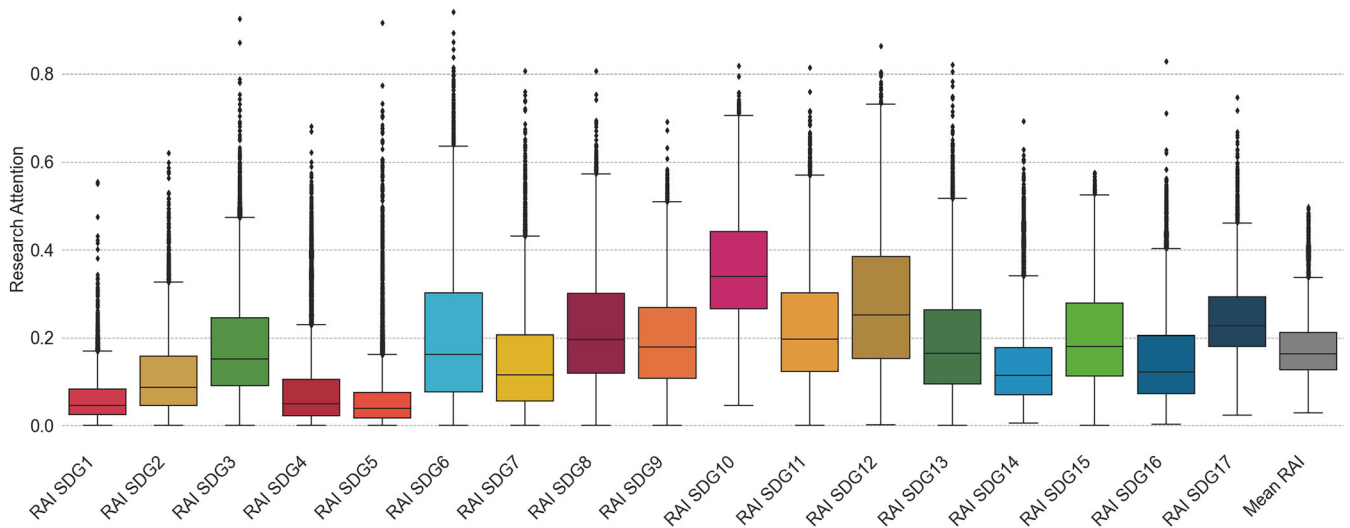


FIGURE A1 Boxplots representing research attention index (RAI) based on the mean of all available abstracts. The figure displays the results of the zero-shot classification at the sentence level for all 13,136 articles. Each sentence in the dataset has been assigned a probability between zero and one for each of the 17 SDGs.

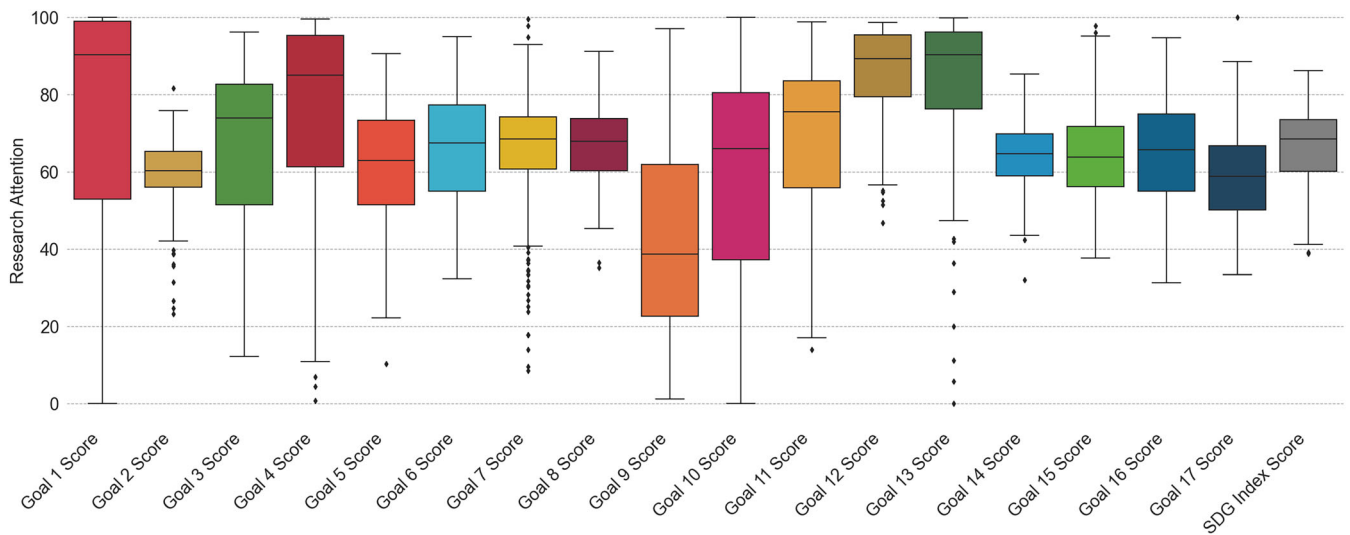


FIGURE A2 Boxplot of official SDG index scores calculated as the mean from 2015 to 2022.

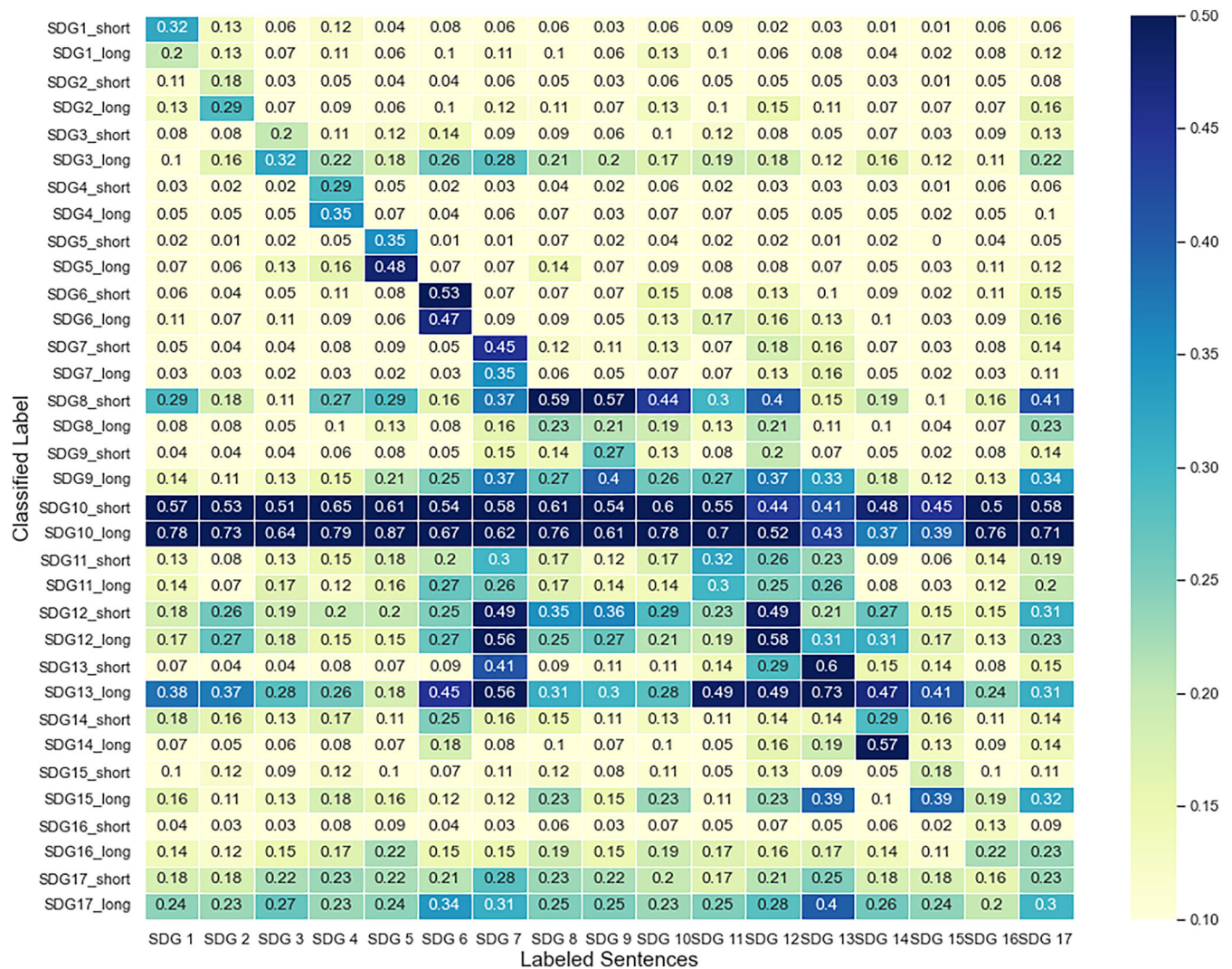


FIGURE A3 Model validation with long and short labels.

TABLE A4 Explanation of labels.

Classified label	Label explanation
SDG1 long	Eradicate extreme poverty
SDG1 short	Extreme poverty
SDG2 long	End hunger, achieve food security and improved nutrition and promote sustainable agriculture
SDG2 short	Eradicate extreme hunger
SDG3 long	Ensure healthy lives and promote well-being for all at all ages
SDG3 short	Good health and well-being
SDG4 long	Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all
SDG4 short	Quality education
SDG5 long	Promote gender equality and empower women
SDG5 short	Gender equality
SDG6 long	Ensure availability and sustainable management of water and sanitation for all
SDG6 short	Clean water and sanitation
SDG7 long	Ensure access to affordable, reliable, sustainable and modern energy for all
SDG7 short	Affordable and clean energy

(Continues)

TABLE A4 (Continued)

Classified label	Label explanation
SDG8 long	Promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all
SDG8 short	Decent work and economic growth
SDG9 long	Build resilient infrastructure, promote inclusive and sustainable industrialization and foster innovation
SDG9 short	Industry, innovation and infrastructure
SDG10 long	Reduce inequality within and among countries
SDG10 short	Reduced inequalities
SDG11 long	Make cities and human settlements inclusive, safe, resilient and sustainable
SDG11 short	Sustainable cities and communities
SDG12 long	Ensure sustainable consumption and production patterns
SDG12 short	Responsible consumption and production
SDG13 long	Take urgent action to combat climate change and its impacts
SDG13 short	Climate action
SDG14 long	Conserve and sustainably use the oceans, seas and marine resources for sustainable development
SDG14 short	Life below water
SDG15 long	Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss
SDG15 short	Life on land
SDG16 long	Promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels
SDG16 short	Peace, justice and institutions
SDG17 long	Strengthen the means of implementation and revitalize the global partnership for sustainable development
SDG17 short	Partnerships for the goals