Response to the Reviewers' Comments (**Reviewer 3**)

Manuscript Title: Smart Homes and Families to Enable Sustainable Societies: A Data-Driven Approach for Multi-Perspective Parameter Discovery using BERT Modelling

Manuscript ID: *sustainability-1882191*

We are very thankful to the reviewer to give us the opportunity to revise our manuscript and for the time and effort in providing very insightful comments and suggestions. These have greatly helped us to improve the quality and presentation of our paper. We have revised the manuscript based on the comments of the reviewers and our response to the specific reviewer's comments is provided below.

All the modifications in the revised manuscript can be viewed using the "Track Changes" option available in MS Word. Our response to the reviewer's comments is provided in the following manner.

> Text written in black colour represents the reviewer's comments.

>>> Text written in blue color represents our response to the reviewer's comments.

We hope that the revised manuscript will meet the reviewer's satisfaction.

>>> Specific Comments and Responses

- This paper introduces a data-driven parameter discovery methodology to provide analysis of the families and homes using over a hundred thousand research papers and nearly a million tweets. A methodology using deep learning, natural language processing (NLP), and big data analytics methods were applied to automatically discover grouped parameters. This is an interesting paper as it has the potential to provide the knowledge and parameters to develop a holistic understanding of matters related to families and homes.

>>> Thank you very much for your positive comments and encouraging remarks. We appreciate it.

Here are some suggestions that might be helpful for the authors to improve their work:

- What does figure 1 serve in this paper? Explanation for figure 1 might be missing.

>>> Figure 1 was explained in Section 1.2 but the reference to Figure 1 was removed by MS Word automatically. We have added the reference to Figure 1.

- Fig. 3 and 4 are not histograms. A histogram should have range value as x-axis and frequency as y-axis.

>>> Thank you for pointing this out. You are right, we wanted to add histograms but somehow it was not coordinated within the team. We have now replaced the earlier figures with the histograms, copied below for your convenience.

Figure 4 shows the histogram for the article abstracts. In this graph, the x-axis provides the word count of the articles in the dataset and the y-axis provides the number of articles in the dataset for a given number of words. Most article abstracts were between 300 and 400 words long. There were a few abstracts that were longer than 600 words.

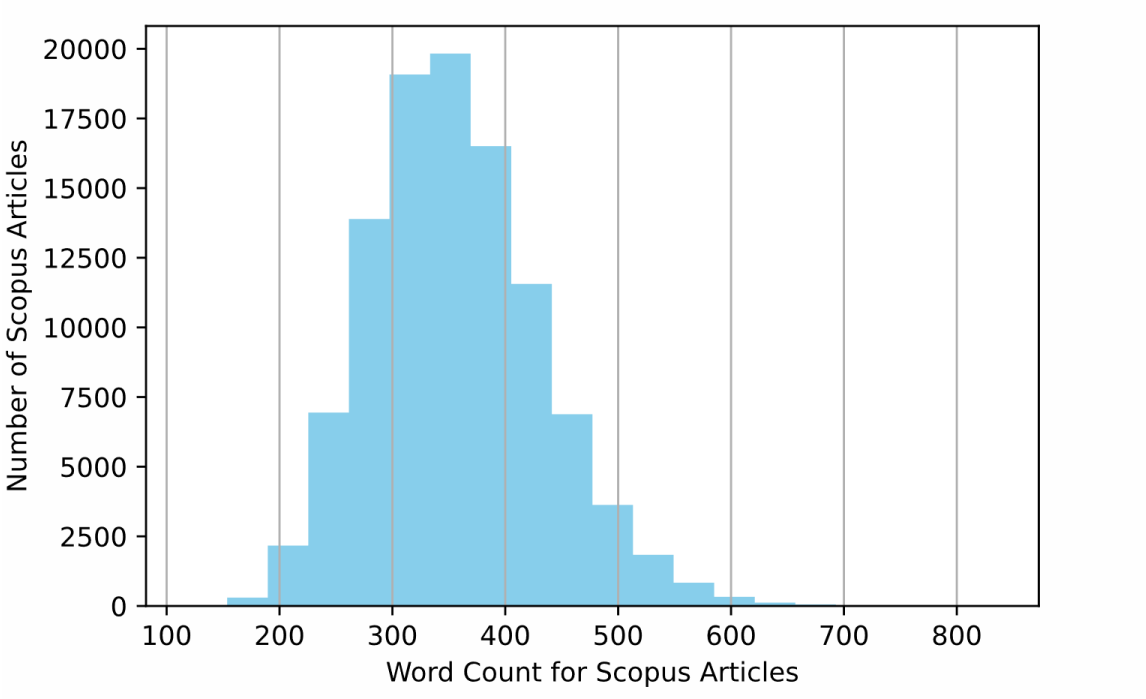


Figure 4 Histogram (Dataset: Scopus).

Figure 5 provides a histogram of the tweets in the dataset. The x-axis shows the number of tweets and the y-axis shows the number of words in the tweets. Approximately 20 to 50 words long are the majority of tweets. There are relatively few tweets that contain more than 75 words. The maximum number of words in a tweet are 120.

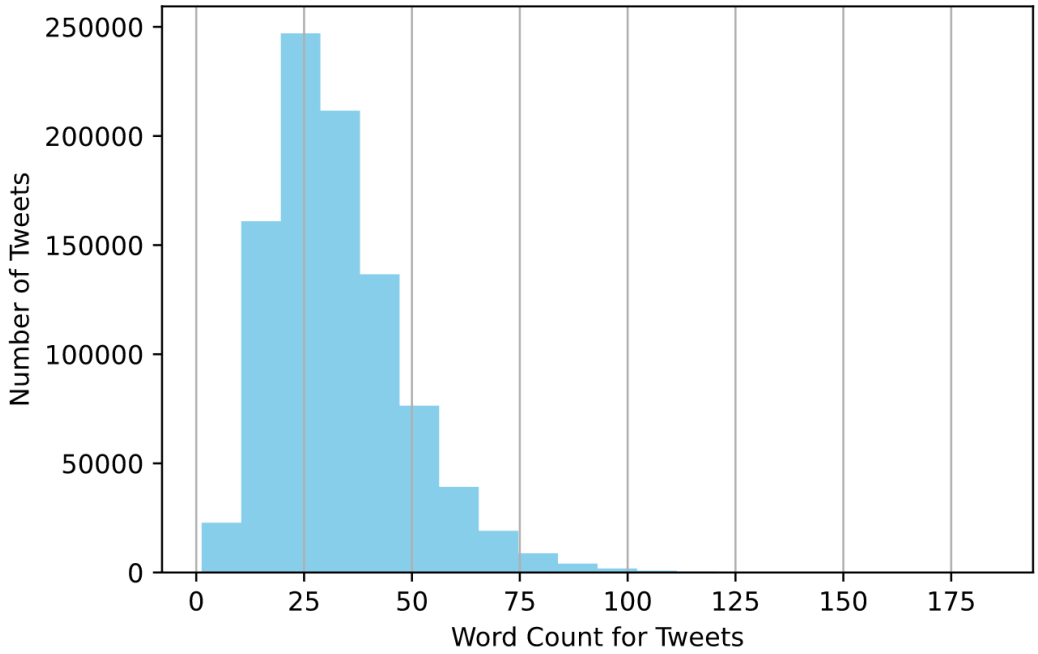


Figure 5 Histogram (Dataset: Twitter).

- The c-TF-IDF, which is defined in equation (1), is essential to understand the following analysis. Therefore, more explanation might be needed for the explanation.

>>> Thank you very much. We have now added additional explanation, copied below for your convenience.

Additionally, the c-TF-IDF score (Class Based Term Frequency-Inverse Document Frequency) is used to calculate words' importance for each document. The TF-IDF score (Term Frequency-Inverse Document Frequency) determines the frequency and prominence of a word in each document and evaluates the word relevance between documents. While the c-TF-IDF considers all the documents of a cluster as a single whole document and then executes the TF-IDF on it, it will achieve a significance score for each word inside a cluster called the c-TF–IDF score. Words that are most significant within a cluster represent that cluster most effectively. In this way, we will be able to obtain keyword-based descriptions for every cluster, which helps to specify the parameter name for it. A c-TF-IDF score is calculated using Equation 1 [34], where f = the word frequency is derived for each class or cluster c and divided by the number of words w. The total number of un-joined documents (d) is then divided by the total frequency of words (f) throughout all classes (cc).

$$c - TF - IDF_C = \frac{f_c}{w_c} \times log \frac{d}{\sum_p^{cc} f_p} \qquad (1)$$

The first part of the equation calculates the Term Frequency (TF) score per class. It can be seen as regularizing frequently used words in a class. The second part determines how common a term is based on its Inverse Document Frequency (IDF) score. The relevance of rare terms is greater than that of common terms. With both TF and IDF scores for each term (word), we can multiply the two to calculate the c-TF-IDF score.

- It is not clear to me how each step, from key words to parameters, and from parameters to Macro are generated. A flowchart showing the entire process might be helpful. For example, how do you decide the keywords under one parameter? And how do you decide what parameters should be in on Macro. You might mention it somewhere in the paper but it would be helpful if you have the flowchart to guide the readers.

>>> Thank you very much. We have added a new paragraph in Section 3.1 to explain the process of parameter discovery, copied below for your convenience. Figure 3 shows the overall methodology.

"The BERTopic extracts a number of parameters from documents without requiring a prior definition of that number. Therefore, it is difficult to predict how many parameters will be extracted from our dataset before training our model. With the knowledge of the number of parameters generated, we specified a reasonable number of parameters using the reduction method, which merges the most similar parameters and then re-calculates the c-TF-IDF to update the representation of our parameters. Finally, the model was saved. Originally, the parameter was represented as an integer number. We used our domain knowledge and quantitative analysis of the keywords and documents in which we examined each parameter's keywords and looked at the context of the keywords in each parameter. This process enabled us to re-label and give appropriate names to each parameter. The process was repeated iteratively to improve parameter names. The process also allowed us to remove irrelevant parameters and merge similar ones. We finally aggregated these parameters into macro-parameters representing higher-level families and home issues."

We hope that the revised manuscript will meet the reviewer's satisfaction.