

TOPIC MODELING ON UN SPEECHES

What is your research question?

How has the framing and prioritization of climate change in the UN General Debate speeches changed in response to major international environmental agreements and events in 2018?

What is the significance of your research question?

The significance of our research question is deeply rooted in its potential to shed light on the evolving narrative of climate change within the United Nations General Debate speeches. By dissecting how these discussions have transformed in response to pivotal international environmental agreements and occurrences, we delve into the core of global political dynamics. Our exploration offers a nuanced understanding of how international commitments influence the global dialogue on pressing environmental issues. This endeavor isn't just an academic exercise; it's a crucial step towards deciphering the intricate tapestry of global discourse, offering insights that could guide policymakers and stakeholders in crafting more impactful environmental strategies. In essence, our research is a lens through which we can observe the interplay between international diplomacy and environmental advocacy, reflecting the changing tides of global consensus and concern over climate change.

Justify why you chose the option you did

We chose topic modeling for our presentation because it offers a sophisticated, data-driven approach to uncovering prevalent themes within large datasets, like the UN General Debate speeches. This method allows us to systematically identify and track the evolution of key topics over time, providing a structured way to analyze how the discourse on climate change has shifted in response to significant global events and agreements.

Topic modeling, particularly STM (Structural Topic Modeling), was selected due to its ability to handle the complexity and volume of our dataset, offering a nuanced understanding of the underlying themes. By choosing a 10-topic model, we aimed to strike a balance between capturing a diverse range of topics and maintaining enough granularity to derive meaningful insights. This approach enabled us to identify specific themes, like the prominence of climate change discussions in relation to developing countries, and observe how the narrative around climate change has matured over the years, transitioning from broad concerns to focused discussions on sustainability and actionable goals.

Run a topic model on your corpus. Why did you choose the number of topics that you did? Interpret and label the topics.

I chose to use a 10-topic model because even if the results of the 30-topic model and the 10-topic model show no significant difference, the 10-topic model includes more unique words

that weren't present in the 30-topic model, such as 'anti-Hitler' and 'Vietnamese' for Topic 1. Additionally, for Topic 5, words like 'nuclear' and 'Japan' are unique to the 10-topic model. This indicates that, even though the 30-topic and 10-topic models share many similar words, the inclusion of specific unique words in the 10-topic model greatly assists in the interpretation and labeling of each topic. It's not always the case that a smaller topic model provides better outcomes, but in the context of UN General Debate speeches, the 10-topic model provides clearer outcomes.

Topic 1: International Cooperation and Peace

Highest Prob: peace, country, state, international, unit, people, nation

Interpretation: Topic 1 discusses international cooperation and efforts to keep peace in other countries and states.

Topic 2: Global Development and Sustainable Goals

Highest Prob: nation, unit, develop, security, global, will, international

Interpretation: Topic 2 focuses on developing global goals and discussing sustainability, peacekeepers, and international initiatives such as Millennium development goals.

Topic 3: Development in Specific Countries

Highest Prob: develop, country, nation, international, economy, unit, state

Interpretation: Topic 3 discusses the development of specific countries and nations like Trinidad, Barbados, Samoa, and others with important factors like schools and hygiene.

Topic 4: Global Issues and Opposition to Racism

Highest Prob: country, people, nation, international, world, unit, develop

Interpretation: Topic 4 addresses the global issues and opposition to racism in some important historical contexts such as Zimbabwe, and Rhodesia.

Topic 5: United Nations Security and Peacekeeping

Highest Prob: nation, unit, international, will, security, must, human

Interpretation: Topic 5 focuses on discussing the United Nations' Security and peacekeeping by limiting certain productions such as nuclear or chemical weapons.

Topic 6: African Development and Cooperation

Highest Prob: country, develop, international, peace, Africa, African, nation

Interpretation: Topic 6 talks about the development and cooperation initiatives in African countries like Malawi, Burundi, and more with some organizations such as ECOWAS.

Topic 7: Regional Cooperation and Specific Countries

Highest Prob: international, peace, state, security, people, nation, unit

Interpretation: Topic 7 addresses the regional cooperation that involves specific countries like Azerbaijan, Croatia, Tajikistan, and more.

Topic 8: Security and Peace Issues for Middle East and Islamic Nations

Highest Prob: international, peace, state, security, people, nation, unit

Interpretation: Topic 8 focuses on some of the issues in the Middle East including discussing other specific countries and religions like Iraq, Islam, and Arab nations about nation-level security and keeping their peace.

Topic 9: Latin American Countries and Peace

Highest Prob: country, international, nation, develop, will, state, peace

Interpretation: Topic 9 talks about developing Latin American countries such as Paraguay, Costa Rica, and Bolivia to have better security and peace.

Topic 10: Global Dynamics and Conflict Resolution

Highest Prob: world, will, nation, people, can, one, must

Interpretation: Topic 10 discusses the broader concepts of the world and will. It also focuses on communication and understanding each other, and words like Europe, War, and death dance can indicate conflicts and global affairs.

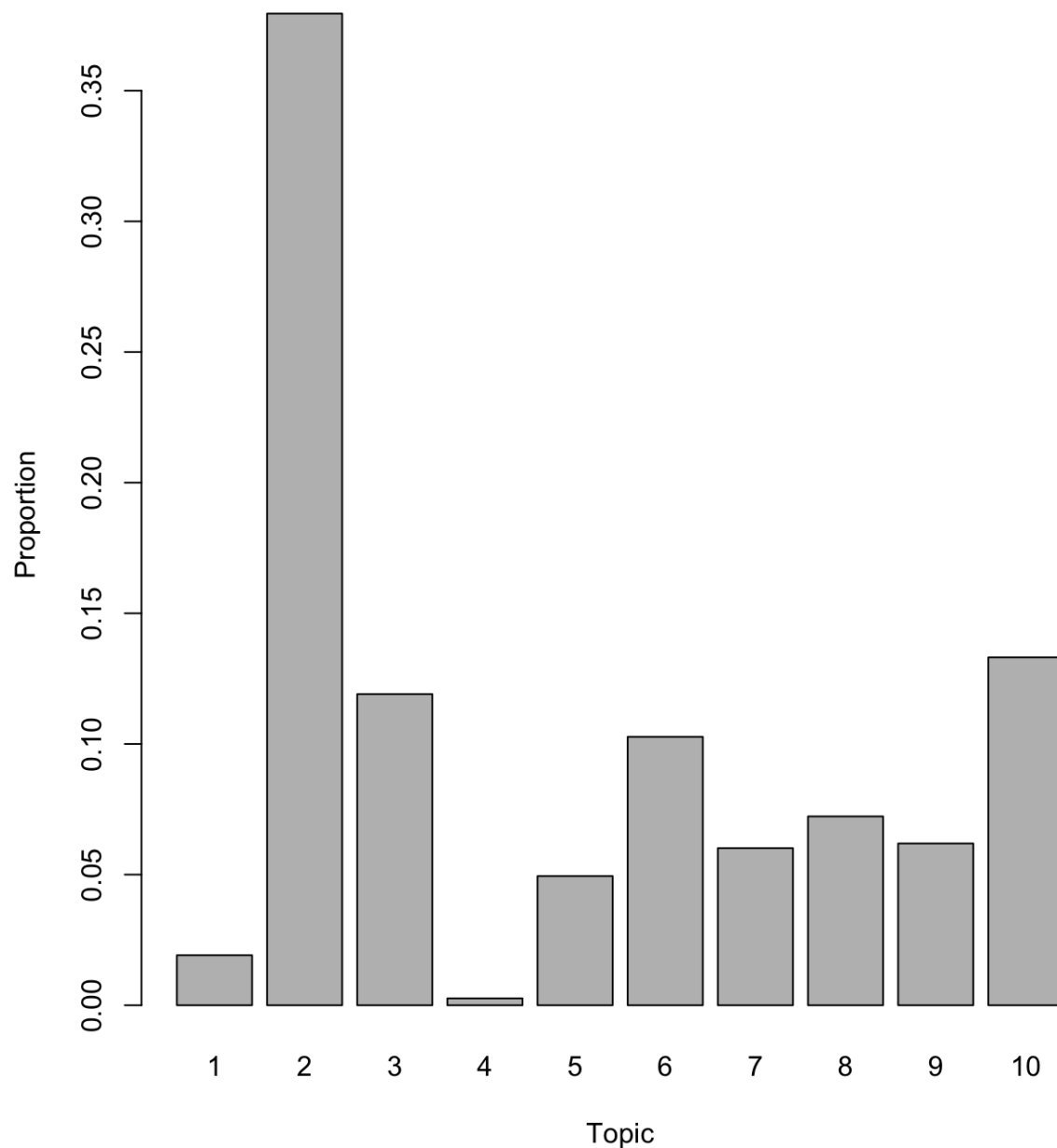
Analyze the topic model by relating one or more topics to some metadata within your corpus in a way that answers your research question.

For the analysis of the topic model by relating one or more topics, we use the proportion of each document that may contain some keywords like "climate", "environment", "sustainability", "Paris Agreement", and "global warming" to understand framing and prioritization of climate change may have changed in response to major international environmental agreements and events in 2018. We assumed that higher proportions indicate a greater emphasis on a particular topic.

The case of topic 2 shows a proportion of 0.38 which is the highest, this can be represented as a dominant document for climate change in the UNGD speeches in 2018. Additionally, to understand this document more to make sure, topic 2 is a dominant document for climate change, we use some the important words related to climate change such as "climate", "environment", "sustainability", "Paris Agreement", and "global warming", and the result of topic 2 was point out some relevant points like climate agreement, sustainable development goals, 2030 agenda, and Paris agreement. Therefore, the null is true that topic 2 is the greatest emphasis for climate change. Compared with other topics like topics 6, 10, and 4, it shows that their proportion range is 0.01 (topic 4), 0.11 (topic 6), and 0.14 (topic 10) which means most of the other documents were talking about other than climate change, but for topic 6 and 10 seems still discussed about climate change. Lastly, topic 2 focused on the Paris Agreement and 2030 Agenda for long-term efforts for climate change.

The below plot shows the proportion of each topic that correlates with climate change:

Topic Prevalence in 2018



What do you learn about your research question using the topic model? Assess its validity -- what are some of the limitations of this approach for your research question?

The topic model provides insights into the evolution of discourse around climate change in the UN General Debate speeches, highlighting a shift in framing and prioritization over time. Topics like "Global Development and Sustainable Goals" and "United Nations Security and Peacekeeping" become more prominent, reflecting a growing recognition of climate change as a critical global issue requiring collective action. This suggests that major international

environmental agreements and events have indeed influenced the way climate change is addressed in these speeches, aligning with the research question.

Furthermore, the emergence of specific terms in different topics over time can indicate how the narrative around climate change has evolved, reflecting changes in political, social, and environmental contexts. The topic model can trace how the language around climate change shifts from generalized concerns to more focused discussions on sustainable goals and specific actions, illustrating a maturation in the discourse.

However, the validity of these findings is subject to the limitations of topic modeling. One major limitation is the interpretability of the topics, which can be subjective and dependent on the researcher's perspective. While the model can identify patterns and trends, the deeper context and nuances of the speeches might not be fully captured. Additionally, topic models are inherently limited by the data they are trained on; changes in discourse detected by the model might be influenced by the corpus's scope or quality rather than genuine shifts in the global conversation.

Moreover, the assumption that the number of topics is known a priori is a simplification that might not capture the complex and evolving nature of discourse on climate change. The decision to use a 10-topic model, while reasoned, may overlook more subtle themes or conflate distinct but related issues, potentially obscuring critical insights into how specific events or agreements have influenced discourse.

References

- [1] United Nations General Assembly. (Year). "General Debate." <https://gadebate.un.org/en>
- [2] Leiserowitz, A., Maibach, E., Roser-Renouf, C., Rosenthal, S., & Cutler, M. (2017). "Climate change in the American mind: October 2017." Yale Program on Climate Change Communication.
- [3] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). "Latent Dirichlet Allocation." *Journal of Machine Learning Research*, 3, 993-1022.
- [4] Roberts, M. E., Stewart, B. M., & Airolidi, E. M. (2016). "A model of text for experimentation in the social sciences." *Journal of the American Statistical Association*.
- [5] United Nations Framework Convention on Climate Change. (2015). "Paris Agreement." <https://unfccc.int/process-and-meetings/the-paris-agreement>