# **REPORT**

### #semanticClimate:

Analyzing Key Structures and Semantic Networks from IPCC Chapter

Submitted by:

SHARON NAYAK

INTERNSHIP

Under the Supervision of

Dr. Gitanjali Yadav Dr. Peter Murray-Rust

NIPGR, New Delhi University of Cambridge,UK



# National Institute of Plant Genome Research New Delhi 110067

December 2024 - March 2025

Table of Contents

**INTRODUCTION3**

**AIMS and OBJECTIVES4**

**DETAILED EXPLANATION5**

**DELIVERABLES10**

**CONCLUSION10**

# Introduction

#semanticClimate is a Science initiative to transform information into structured, filtered, and actionable knowledge. It liberates knowledge from climate-related reports.

Goals:-

* To convert the IPCC documents from HTML to XML
* Extract terms and explore their use and meaning
* Link terms and explore their use and meaning.
* Create new structures for navigation, search, display.

The United Nations periodically publishes climate reports through several channels including the Intergovernmental Panel on Climate Change (IPCC) and United Nations Framework Convention on Climate Change (UNFCCC). It is important to realise that the availability of these reports to the public does not ensure their accessibility.

The focus is on two main activities:

* developing software tools for searching climate change literature, and
* fostering community engagement through events like hackathons, datathons, data workshops and panel discussions.

The toolkit developed by #semanticClimate uses Artificial Intelligence (AI) over Natural Language Processing (NLP) to transform locked literature (PDFs) into semantic hypermedia (HTML/XML).

The software tools includes:-

* *pygetpapers*, which helps users find and download scientific articles, and
* *docanalysis*, which analyses these articles for important terms and information.

These tools also support image processing, document conversion, and the creation of knowledge graphs and translations using resources like Wiki data and Wikimedia.

At present, our latest tools *amilib* and *amiclimate* are transitioning from their alpha to beta phase, which means the software is becoming ready for public use. This shift involves scaling community activities and adapting to new challenges, such as exploring UN Climate corpora and expanding citizen science events across India. We believe that openness is a powerful accelerant for finding solutions to our most pressing challenges.

# Aims and Objectives

The IPCC prepares comprehensive Assessment Reports about knowledge on climate change, its causes, potential impacts and response options. It also produces Special Reports, which are an assessment on a specific issue and Methodology Reports, which provide practical guidelines for the preparation of greenhouse gas inventories.

Materials are published generally in three volumes, one for each of the Working Groups of the IPCC, plus a Synthesis Report. Each of the Working Group volumes is composed of individual chapters, an optional Technical Summary and a Summary for Policymakers.

* Working Group I examines the physical science underpinning past, present, and future climate change.
* Working Group II assesses the impacts, adaptation and vulnerabilities related to climate change.
* Working Group III focuses on climate change mitigation which involves actions that reduce the rate of climate change.

**As a part of #semanticClimate, my tasks during the internship tenure included:**

1. Exploring the toolkit used by the team, and understanding the cause that #semanticClimate is working towards. Going through the provided resources and documentation of the software developed in-house.

2. Participated in Workshop:- “Hands-on Workshop on BioAI for Research”

3. Overviewing the chapters of the Sixth Assessment Report (AR6) and selecting one to contribute to, by creating a wordlist for the comprehensive semantic Dictionary.

I undertook work on Cross-Chapter 7, of Working Group II - “**TROPICAL FORESTS**”.

# Detailed Explanation

In this section, all of my detailed intern work is listed:-

1. **Summary of my cross-chapter : - “TROPICAL FORESTS”**

*Over 420 million ha of forest were lost to deforestation from 1990 to 2020. More than 90% of that loss took place in tropical areas, threatening biodiversity, environmental services, livelihoods and resilience to climate shocks. Despite CO2 fertilisation, ongoing climate change has weakened the carbon sink potential of tropical forests in Amazonia and, to a lesser extent, in Africa and Asia. Community-based adaptation, built on Indigenous knowledge and local knowledge (IK and LK) over centuries or millennia, is often identified as an effective adaptation strategy to climate change. Forest restoration using a diverse mix of native species can help rebuild the climate resilience of tropical forest ecosystems, but is best implemented alongside other sustainable forest management strategies and adaptation interventions (high confidence). The urgency of tackling drivers of forest loss and degradation is increasing as climate impacts on forests and ecosystems increase (high confidence), the authors say. The authors conclude that the world needs to take action to reduce deforestation and forest degradation in the tropics and around the world in the next few decades. They call for policies that create incentives for environmental services such as carbon storage and biodiversity refugia, as well as for the creation of ‘forest reserves’ to store carbon in the form of trees, grasses and shrubs. The report is published by the Royal Society of Tropical Biology and Climate Change (RCBC) on behalf of the World Resources Institute (WRCI) (http://www.wrci.org/content/2013/07/28/wrcitl/climate-change-and-forest-resilience-in-the-tropics.html#storylink=cpy).In Europe, the European Union (EU) and the UK (UK), the European Commission (UK) are also working on a range of issues relating to deforestation, climate change and forest resilience, including climate change adaptation and mitigation, such as the Paris Agreement on Climate Change and the European Convention on Biological Diversity (ECONOMIC). In the Middle East and North Africa, the EU is working on an action plan to tackle deforestation in the region, including a new strategy to tackle climate change in the South East.*

This summary is extracted by using HuggingFace Tools, a Transformer model : - T5 and BART Model for summarization.

1. **Wordlist of Cross-Chapter**

It is made by using *nltk.* NLTK, or the Natural Language Toolkit, is a widely used, open-source Python library for working with human language data. Thus, it helps to get words from the cross-chapter.

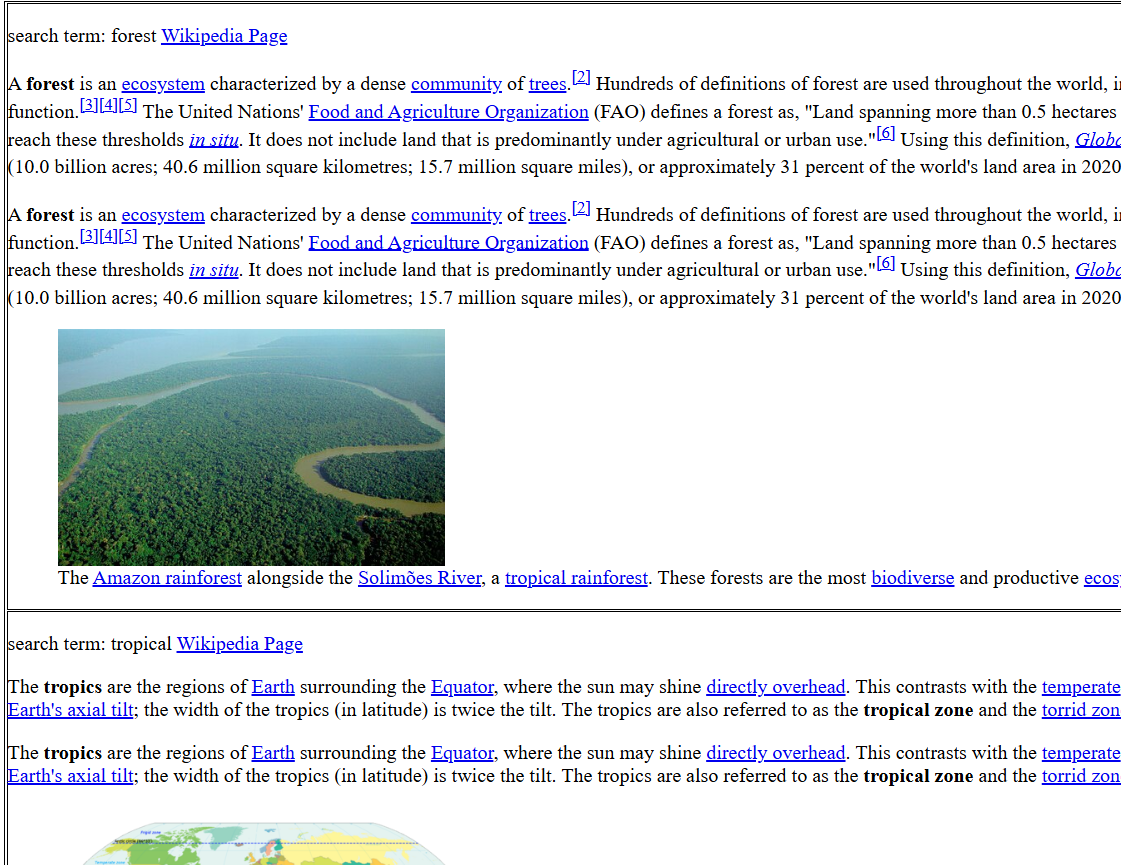


1. Dictionary of Cross-Chapter

Dictionary is made with the help of the *amilib* tool.

Creating a Wikimedia [enhanced dictionary with *amilib* from wordlist](https://colab.research.google.com/drive/1QNETQ3bZFgOvu2iyZCZ0jM9tjTWuUiPi?usp=sharing).

A google colab has been made in which *amilib* is imported and used to make a dictionary by help of Wikimedia definition.

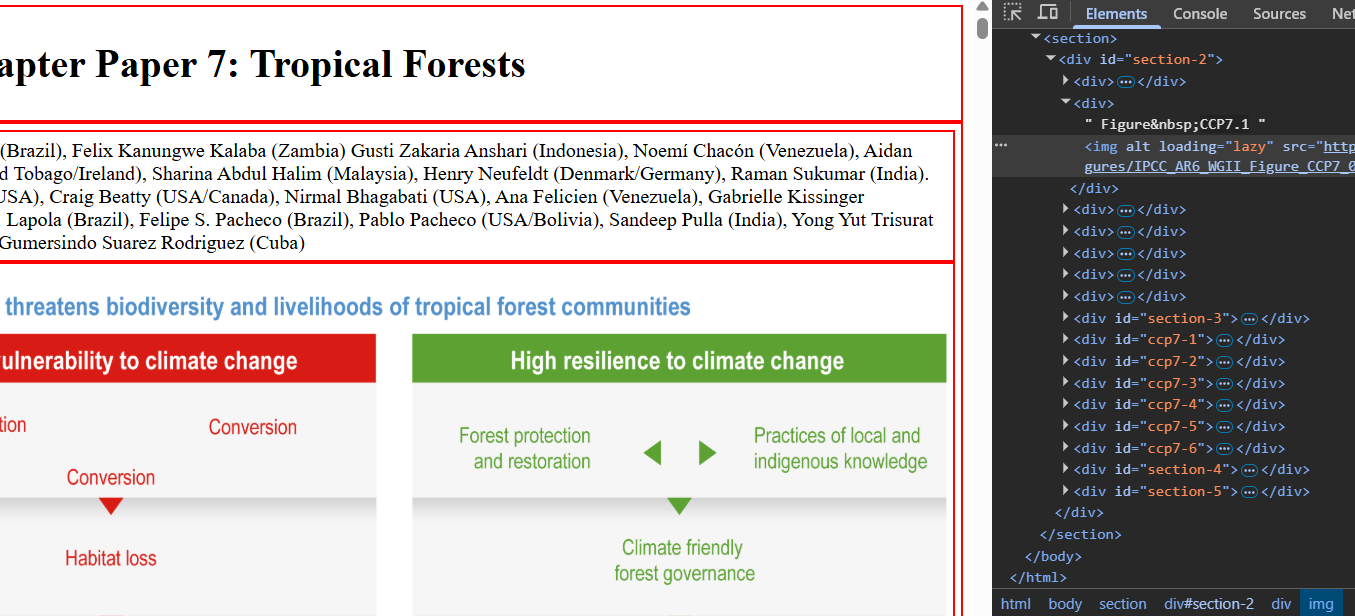


1. **Created *html\_with\_ids* for the cross-chapter**

To make the graph of the cross-chapter, *html\_with\_ids* for the chapter is required.  
This has been created with the workflow approach: -

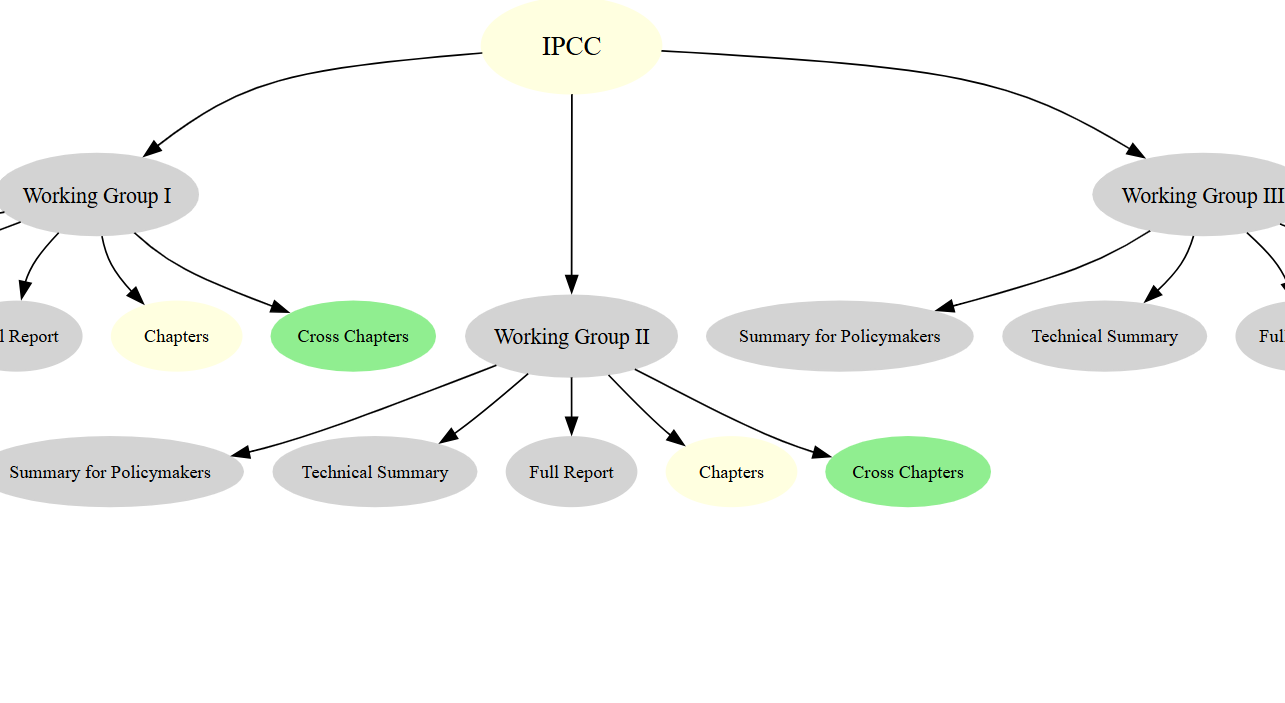
1. *Fetch and clean the HTML.*
2. *Load deletion commands from the JSON file.*
3. *Strip and modify the HTML based on the commands.*
4. *Save the final stripped HTML to the output file.*

Output file consists of divs which have ids associated with it.



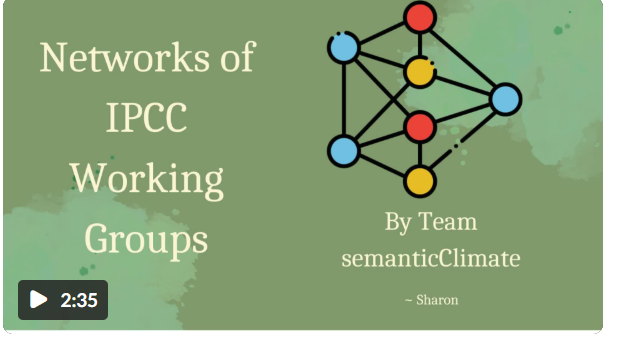
5. **IPCC Graph Structure**

Graph has been made for easy access of IPCC reports and chapters. It has been made with the help of the Graphviz package in the Python Language.



6**. Created a Video on NetworkX**

A simple video on NetworkX for Brussels Workshop.



A screenshot of a computer

AI-generated content may be incorrect.

## Deliverables

* Link for the summary: -<https://github.com/semanticClimate/internship_sC/blob/sharon/CC7summarization.ipynb>
* Link for the wordlist : - <https://github.com/semanticClimate/internship_sC/tree/sharon/Wordlist>
* Link for the dictionary: - <https://github.com/semanticClimate/internship_sC/blob/sharon/CCP7Dictionary.ipynb>
* Link for the IPCC graph: -<https://github.com/semanticClimate/internship_sC/blob/sharon/ipcc_structure.svg>
* Link for all the work in SemanticClimate: - <https://github.com/semanticClimate/internship_sC/tree/sharon>

## Conclusions

In conclusion, my experience at SemanticClimate was both enriching and transformative. Over the course of my time there, I developed a strong understanding of professional work environments, including essential practices like effective teamwork, meeting strict deadlines, participating in daily stand-up meetings, and confidently delivering presentations. These experiences significantly improved my communication and collaboration skills.

On the technical front, I was introduced to several important tools and methodologies. I learned to use Git for version control, which taught me the importance of maintaining clean and organized code. I also explored HTML page processing, built network graphs to visualize data relationships, and applied advanced summarization techniques using HuggingFace. These experiences have deepened my interest in data processing and AI-powered tools.

Overall, my time at SemanticClimate helped me grow both personally and professionally, and I look forward to applying these skills in future projects.