Building a System Model on The Interlinkages of Climate Action with Sustainable Development Goals (SDGs): Using natural language processing for systematic text mining of the key linkages from climate change literature

Xin Zhou[1], Kshitij Jain[2], Mustafa Moinuddin[1] & Patrick McSharry [3,4,5]

[1] Institute for Global Environmental Strategies (IGES)
[2] Google, Canada
[3] Carnegie Mellon University
[4] Kigali Collaborative Research Centre
[5] Oxford University

International Conference on SYSTEMS ANALYSIS FOR ENABLING INTEGRATED POLICY MAKING
August 10-12, 2022, Scope Convention Centre, New Delhi

The United Nations' Sustainable Development Goals (SDG) with 17 Goals and 169 Targets form an interacted and indivisible framework

Importance of taking an integrated approach to achieve the SDGs through their intrinsic interlinkages

- Taking an integrated approach is imperative for achieving SDGs as a whole.

- Such an integrated approach is new and challenging.
  - Broad coverage of social, economic and environmental dimensions;
  - Complicated interactions among 169 targets.

- Scientific knowledge on the interlinkages between climate action and SDGs is limited which inhibits practising integrated policy making.


Interconnected SDGs and targets requires a holistic approach to strength the synergies and mitigate the trade-offs.

Emerging attempts to understand and analyse the linkages among the SDGs

- **Different approaches**
  - Expert opinion (e.g., ICSU 2017),
  - Literature review (e.g., Zhou et al., 2022)
  - Statistical analysis (e.g., Dörgő et al., 2018).

- **IGES' SDG Interlinkages Analysis & Visualisation Tool**
  - Identify, quantify and visualise the interlinkages between the SDG Targets.
  - Causal links identified through a literature review conducted manually
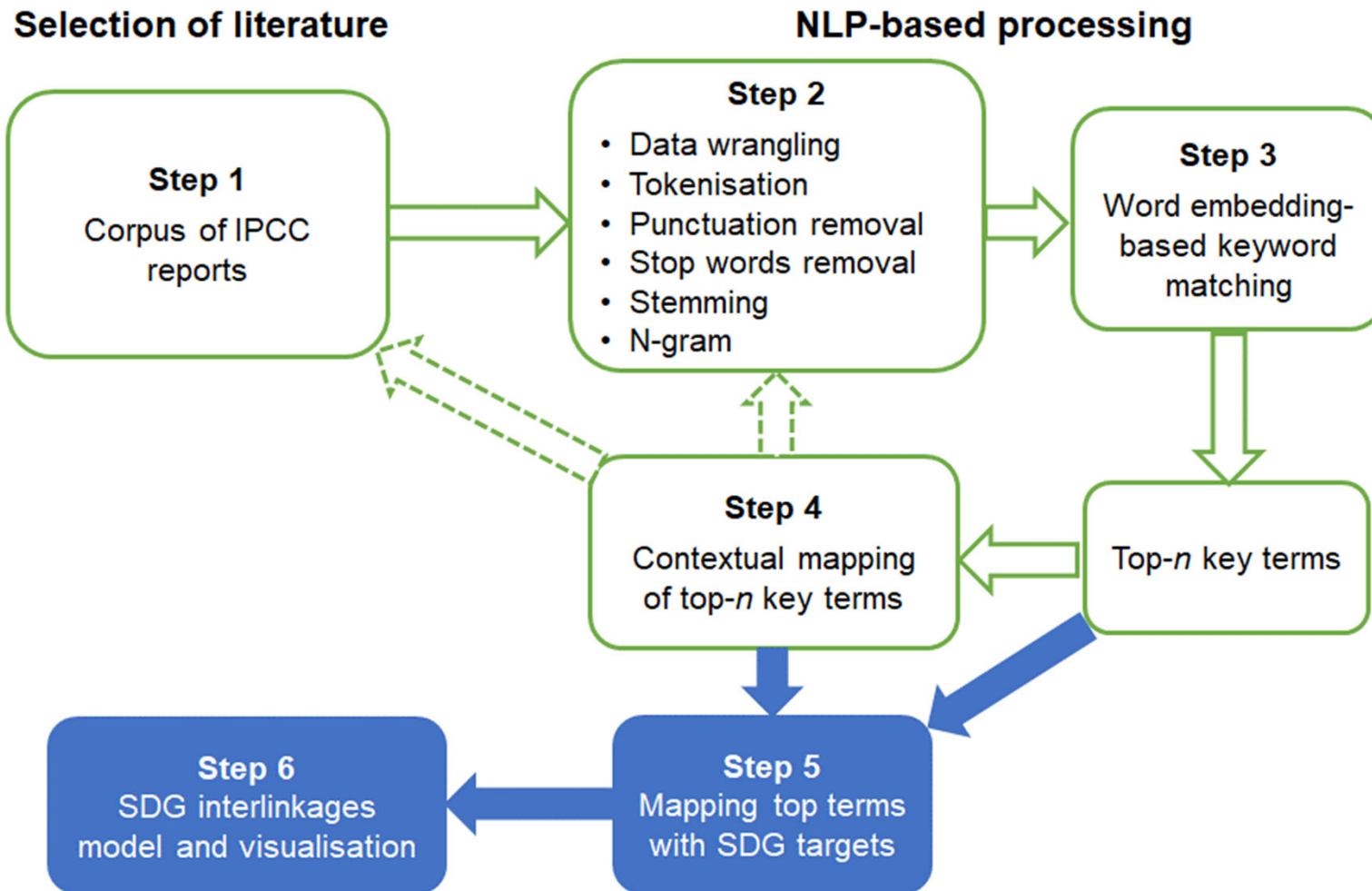

https://sdginterlinkages.iges.jp/visualisationtool.html

4
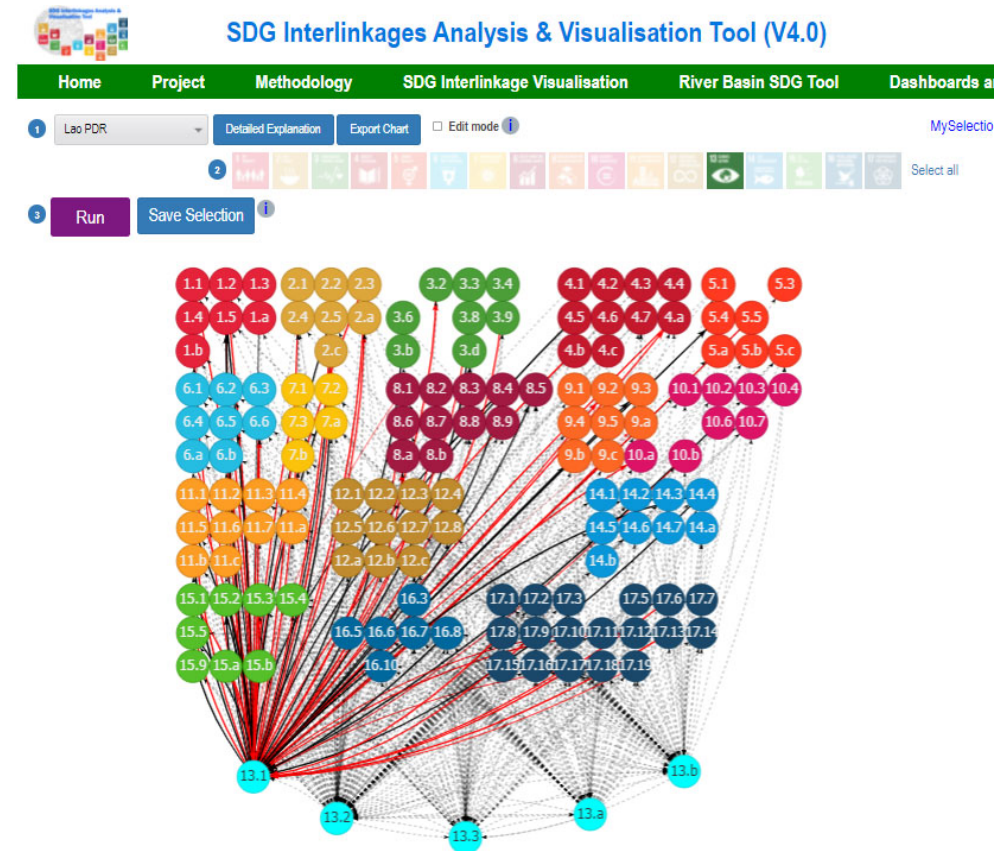
Limitations of the existing method and the purpose of this study

- **A literature review conducted manually is limited**

  - laborious and time-consuming

  - subjective based on experts' judgement

  - unable to capture a full spectrum of the interactions in a systematic way

- **An AI-based systematic and automated approach is needed**

  - To help systematically process a massive quantity of literature

- This study aims at developing an automated process by using Natural Language Processing (NLP) to systematically extract the SDG linkages from climate change literature and build a system model on the interlinkages of climate action with SDGs.

Methodology: Using NLP to systematically extract the key SDG interlinkages from climate change literature

Selection of literature

NLP-based processing

Step 1
Corpus of IPCC reports

Step 2
- Data wrangling
- Tokenisation
- Punctuation removal
- Stop words removal
- Stemming
- N-gram

Step 3
Word embedding-based keyword matching

Step 4
Contextual mapping of top-*n* key terms

Top-*n* key terms

Step 5
Mapping top terms with SDG targets

Step 6
SDG interlinkages model and visualisation

Using NLP for extracting top 500 terms

🌀 By running a hyper-parameter search on the set of n-grams from the cleaned tokens, 2, and 3 n-grams are found best suited for our purpose. We selected top 500 bigrams and trigrams by their frequency.

Top 500 key terms (part of the results)	Total frequency
climate change, climate changes, climatic changes, climate chang, climate changed, climatic change, climate changing, climates change, climat change, climates changes, climate changs, climates changing	43548
developing countries, developed countries, developing country, developed country, development countries, development country, develop countries	5684
seal level, seal levels	4770
greenhouse gas, greenhouses gas	3883
impacts climate, impact climate, impacts climatic, impact climatic, impacted climatic, impact climat, impacts climat, impacted climate, impacting climate	3840
impacts climate change, impact climate change, impact climate changes, impacts climatic change, impact climatic changes, impacts climatic changes, impact climatic change, impacts climate changes, impacted climatic changes, impact climat change, impacts climat change	3357
energy use, energy used, energy uses, energy useful, energies use, energy using	1699
fossil fuels, fossil fuel	2337
land use, land uses, land used, lands use, lands using, lands used, land using	2604
energy efficiency, energy efficient, energy efficiencies, energy efficiently, energy efficie	2206
water resources, water resource, water resourc	2099
climate policies, climate policy, climatic policy	1997
adaptive capacity, adaptation capacity, adapt capacity, adaptive capacities, adaptation capacities, adapt capacities, adaptative capacity	1529
extreme events, extreme event, extremes events, extremal events	1374
climate change mitigate, climate change mitigation, climate change mitigated, climate change mitig, climatic change mitigation, climate change mitigates	1254
renewable energy, renewable energies, renewables energy, renew energy, renewal energy	1340
human health	1016

Using NLP for extracting the contextual terms of the top terms and the description of their relationships in the contextual sentences
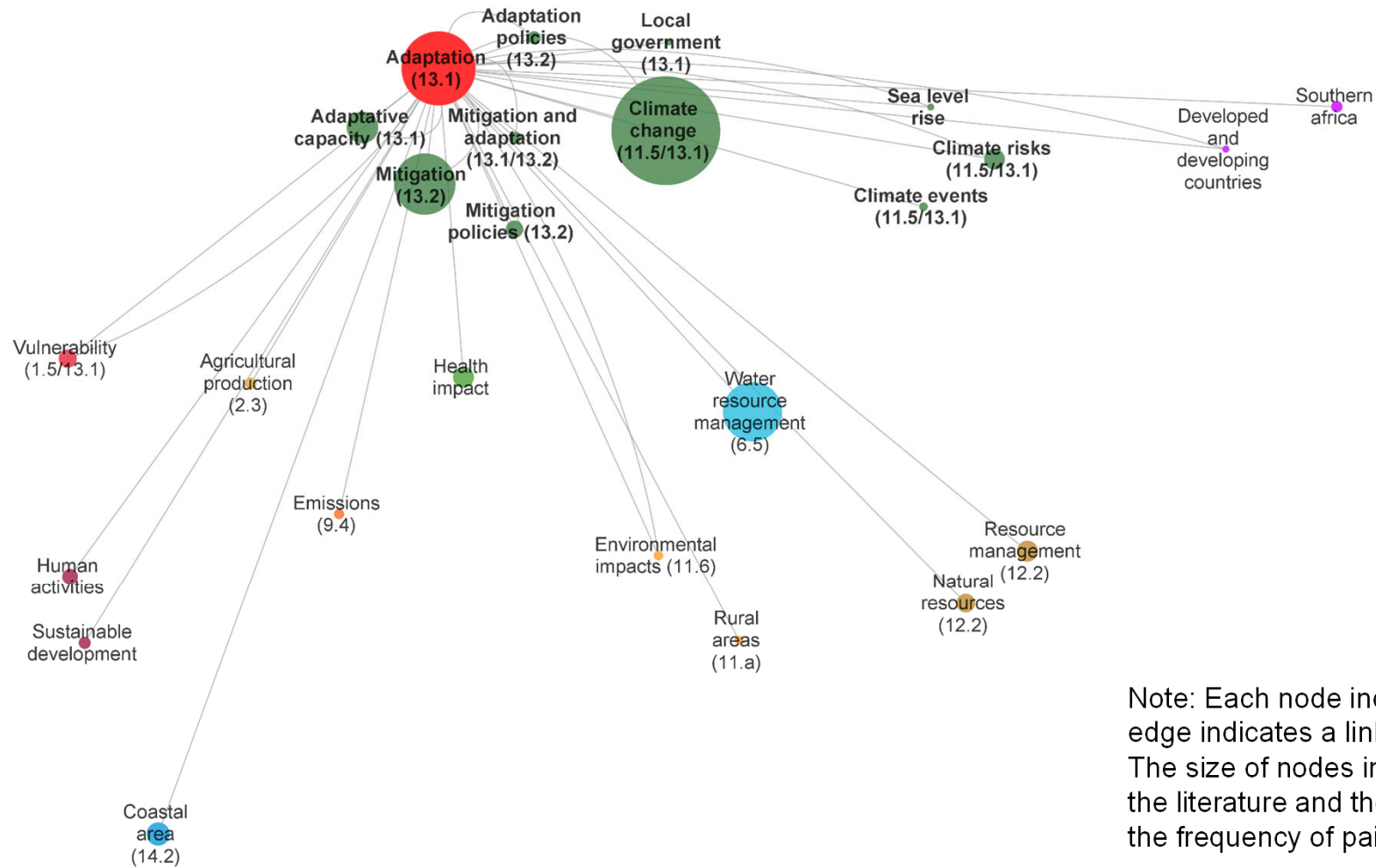
- For each top term, we extracted 10 key contextual terms for each of the top 500 key terms based on frequency, together with the contextual sentence describing the relationships of the pair terms. 4,600 pairs were identified.

Center Terms	Context Terms	Total Score	Context Sentenece	Pre or Post
impacts climate, impact climate, impacts climatic, impact climatic, impacted climatic, impacted climate, impacting climate	water resources, water resource, water resourc	43	['impacts', 'climate', 'change', 'hydrology', 'water', 'resources']	post
			['water', 'resources', 'since', 'publication', 'ipcc', 'first', 'impacts', 'number', 'impacts', 'climate']	pre
			['impacts', 'climate', 'change', 'hydrology', 'water', 'resources']	post
			['water', 'resources', 'socioeconomic', 'impacts', 'etc', 'impacts', 'climate']	pre
			['impacts', 'climate', 'change', 'natural', 'terrestrial', 'ecosystems', 'socioeconomic', 'consequences', 'impacts', 'climate', 'change', 'hydrology', 'water', 'resources']	post
impacts climate, impact climate, impacts climatic, impact climatic, impacted climatic, impacted climate, impacting climate	human health	27	['impacts', 'climate', 'change', 'human', 'settlement', 'energy', 'transport', 'industrial', 'sectors', 'human', 'health']	post
			['human', 'health', 'air', 'quality', 'impacts', 'climate']	pre
			['impacts', 'climate', 'change', 'human', 'settlement', 'energy', 'transport', 'industrial', 'sectors', 'human', 'health']	post
			['impacts', 'climate', 'change', 'human', 'settlement', 'energy', 'transport', 'industry', 'human', 'health']	post
			['impacts', 'climate', 'change', 'human', 'health']	post
water resources, water resource, water resourc	water supply, water supplies, water supplied	40	['water', 'resource', 'systems', 'less', 'hydrometeorological', 'changes', 'assumed', 'addition', 'changes', 'water', 'supply']	post
			['water', 'resource', 'systems', 'less', 'hydrometeorological', 'changes', 'assumed', 'greenhouse', 'addition', 'changes', 'water', 'supply']	post
			['water', 'resource', 'management', 'systems', 'producing', 'reasonably', 'accurate', 'water', 'supply']	post
			['water', 'resource', 'systems', 'less', 'hydrometeorological', 'changes', 'assumed', 'greenhouse', 'addition', 'changes', 'water', 'supply']	post
			['water', 'supply', 'billions', 'gallons', 'day', 'current', 'hypothetical', 'climatic', 'conditions', 'major', 'water', 'resource']	pre

A qualitative SDG interlinkages model for Goal 13 on climate action

Note: The figure was generated by using Cytoscape. Each node indicates a top term and the size of nodes indicates their frequency. The code in parentheses indicates the corresponding SDG targets. The edge indicates a linkage between paired terms. The width of an edge indicates the frequency of the paired terms.

An example of adaptation and linkages with SDGs

Note: Each node indicates a key term and each edge indicates a link between a paired terms. The size of nodes indicates their frequency in the literature and the width of edges indicates the frequency of paired terms.

Value-added, areas of application and future research agenda

- By using advanced NLP techniques, an automated process for systematically extracting the key terms and their contextual terms was developed and a qualitative system model on the interlinkages between climate action and SDGs was built.

- **Areas of application**
  - The model can help inform policy makers on the interactions between climate action and SDGs and support climate action plan development;
  - The model can be used to communicate among stakeholders and raise awareness on the impacts of climate change and raise urgency for taking actions.
  - The methodology can be extended to all 17 Goals and updating the existing IGES Interlinkages Tool to support integrated planning and priority setting.

- **Future research agenda**
  - Building a methodology for selecting relevant and quality scientific literature from bibliographic databases (Google Scholar, SCOPUS, etc.);
  - Improving NLP-based text data processing, such as excluding irrelevant n-grams and including all relevant n-grams supported by PoS tagging, identifying synonyms, and defining causations, etc.

Thank you very much for you kind attention!

Contact:
zhou@iges.or.jp
kshitijj@google.com
moinuddin@iges.or.jp
mcsharry@cmu.edu