



**Hertie School**

Data Science Lab



# Using NLP to Track Health Implications of Climate Change

Slava Jankin

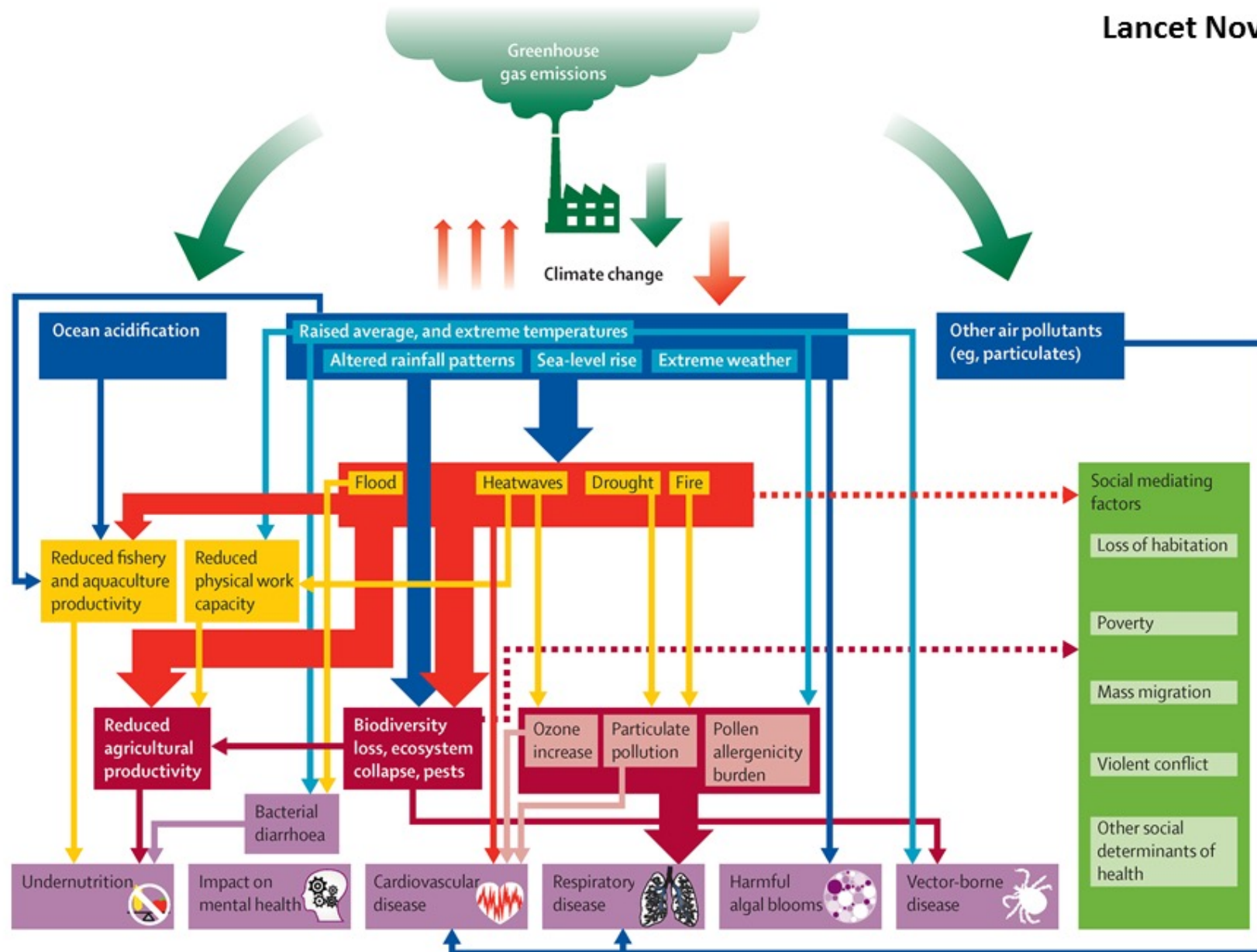
Data Science Lab

Hertie School, Berlin

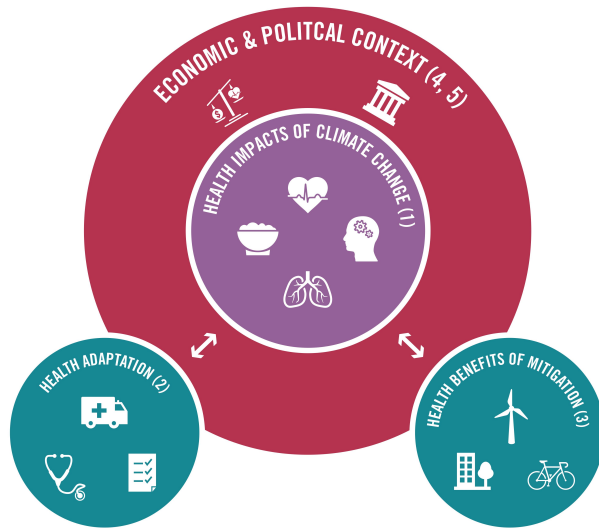
1st Workshop on Computational Linguistics for Political Text Analysis  
KONVENS 2021

# Health effects of global climate change

Lancet Nov 2015

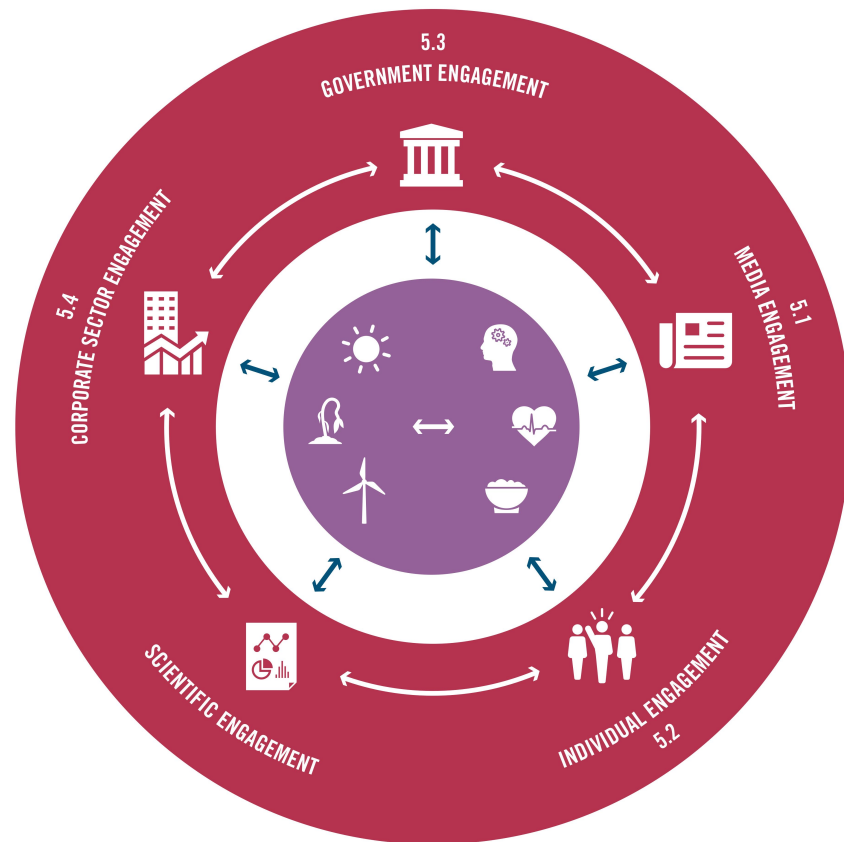


# The Lancet Countdown



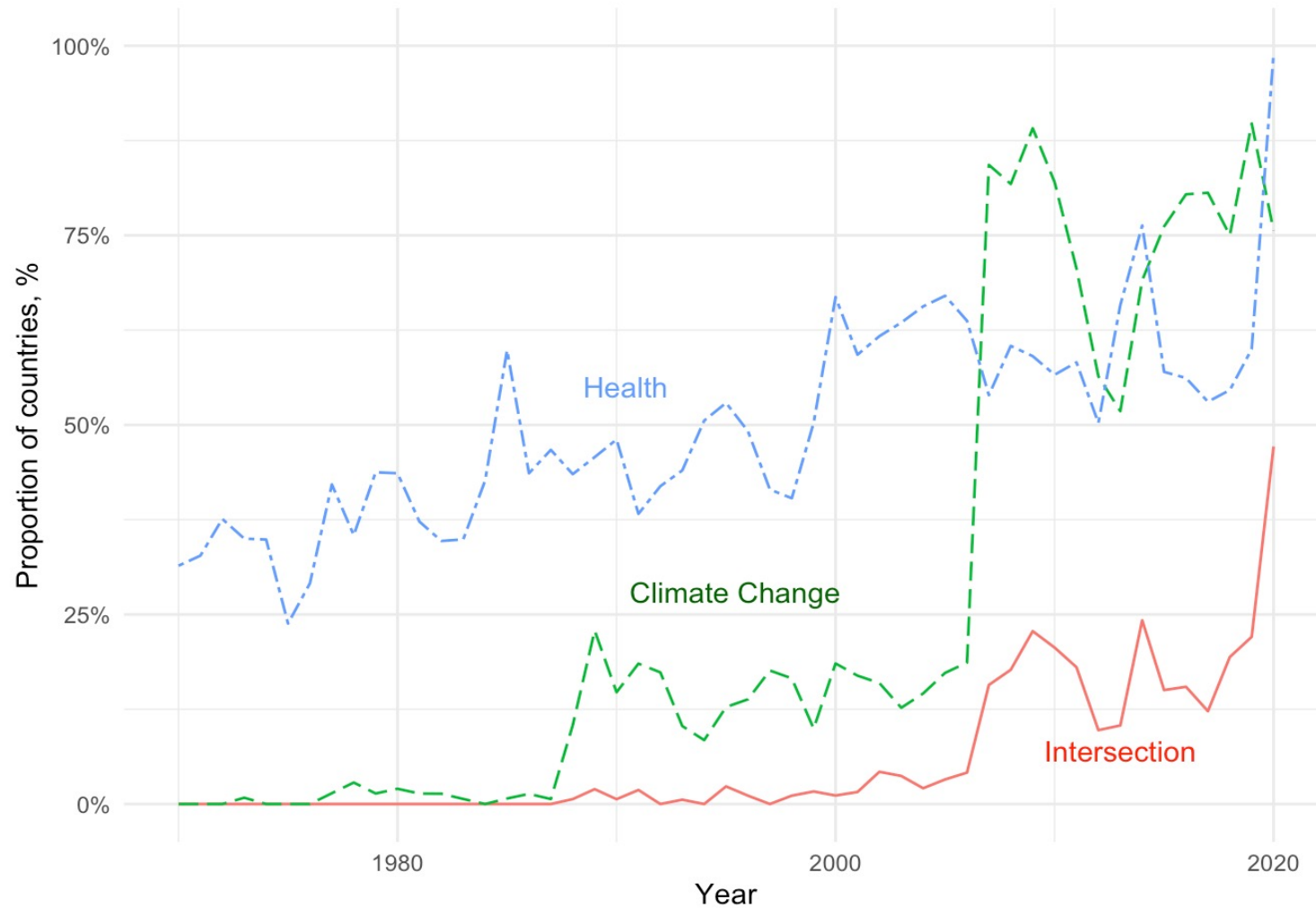
- 35 academic and UN institutions
- 120 scientists

# Public and Political Engagement



- 5.1 – Media Coverage Of Health And Climate Change
- 5.2 – Individual Engagement in Health and Climate Change
- 5.3 – Engagement in Health and Climate Change in the United Nations General Assembly
- 5.4 – Engagement in Health and Climate Change in the Corporate Sector

# 5.3 Government Engagement with Health and Climate Change



# Building the indicator



Indicator must have:

- (absolute) global coverage
- regularity
- meaningfulness
- resource constraint (pro bono)

United Nations General  
Debate



# Building the indicator: data



## UN General Debate Corpus

- All speeches from 1946 to 2019.
- UN membership: 55 (1946), 70 (1970) and 193 (2019).
- United Nations Dag Hammarskjold Library
- Official English language version.
- ~10K documents.
- On average 123 sentences per speech, 945 types, and 3,248 tokens each.
- Harvard Dataverse repo



# Building the indicator: data cleaning

JB/12

A/44/PV.4

46

The meeting was suspended at 11.25 a.m. and resumed at 11.50 a.m.

ADDRESS BY MR. GEORGE BUSH, PRESIDENT OF THE UNITED STATES OF AMERICA

The PRESIDENT: The Assembly will now hear an address by the President of the United States of America.

Mr. George Bush, President of the United States of America, was escorted into the General Assembly Hall.

The PRESIDENT: On behalf of the General Assembly I have the honour to welcome to the United Nations the President of the United States of America, His Excellency Mr. George Bush, and to invite him to address the General Assembly.

President BUSH: I am honoured to address the General Assembly today at the beginning of its forty-fourth session. I should like to congratulate Joseph Garba of Nigeria, a distinguished diplomat, on his election as President of this session of the General Assembly, and I wish him success in his presidency.

I feel a great personal pleasure on this occasion, for this is a homecoming

# Building the indicator: validating data

- “Speeches at the General Debate are interesting because they flesh out national policies; what states think. It is the one place where states can speak their mind – it reflects the issues that states consider to be most important. In this way, the speeches in the General Debate are unique... The [GD] speeches are one of the least coordinated parts of the [EU’s] UN activities.” –Deputy Head of Finnish Mission to the UN
- “coordination for the General Debate does not happen; governments do not want other states looking over their speeches ... these speeches are the most sovereign thing that a country does as a member of the UN.” –Spokesperson for the German Mission to the UN
- “the General Debate is important because it officially sums up the position of Italy on the most important things in international politics in that particular year... the speech is a summary of a country’s foreign policy in a given year.” –Diplomat for the Italian Mission to the UN

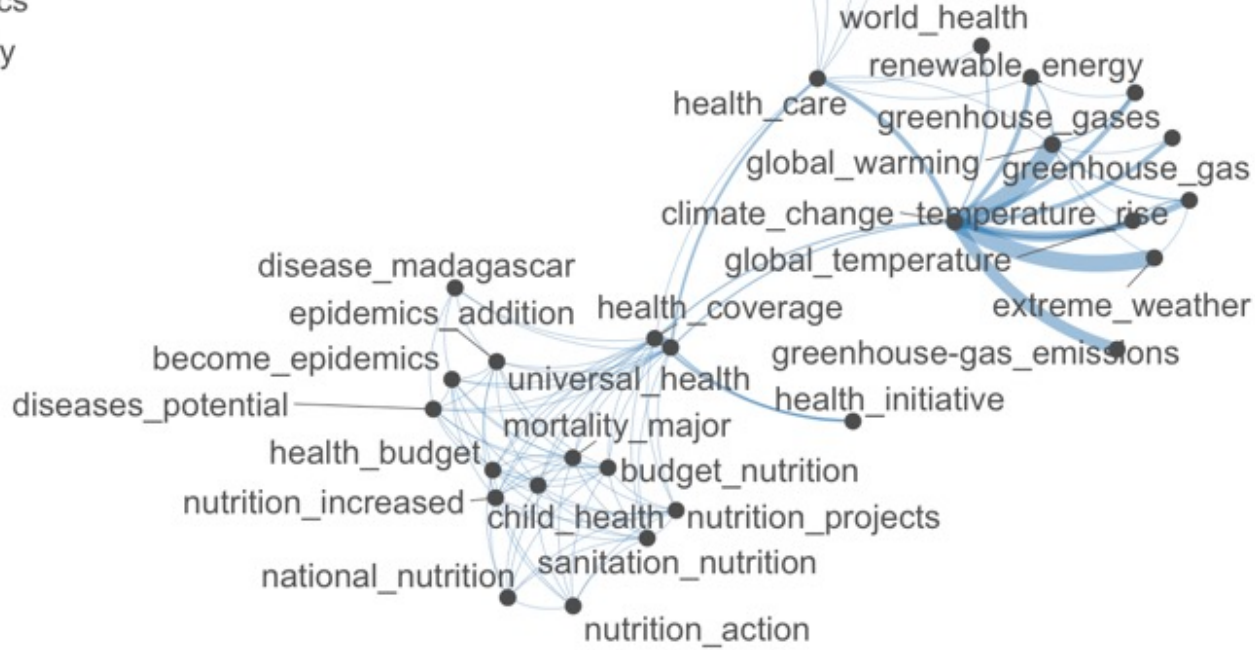
# Building the indicator: methods

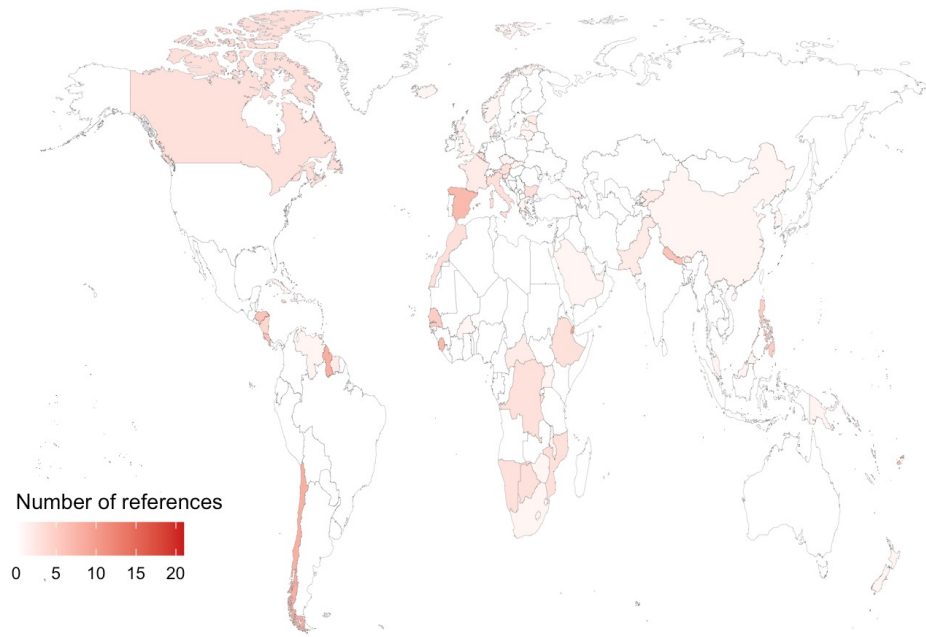
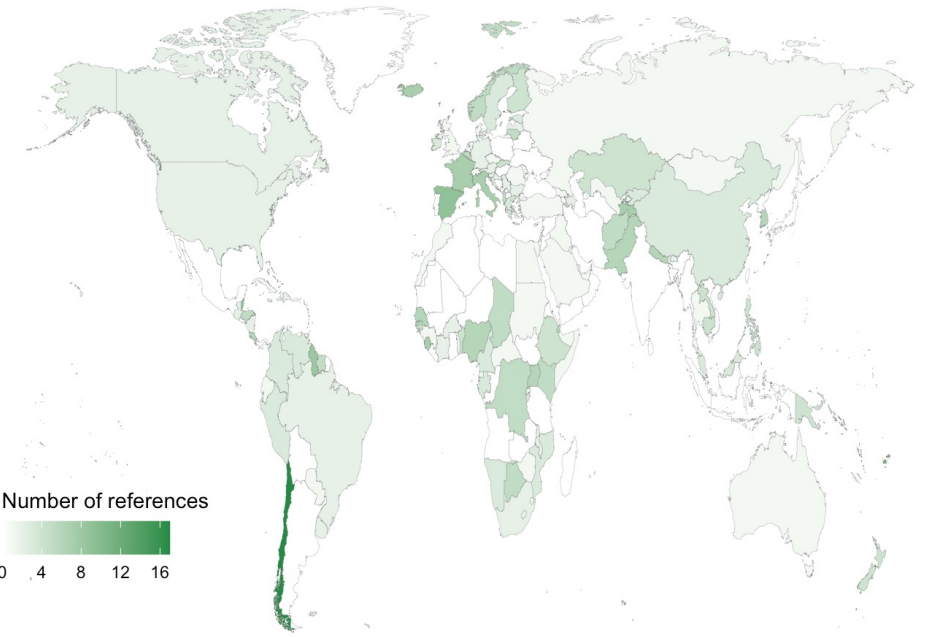
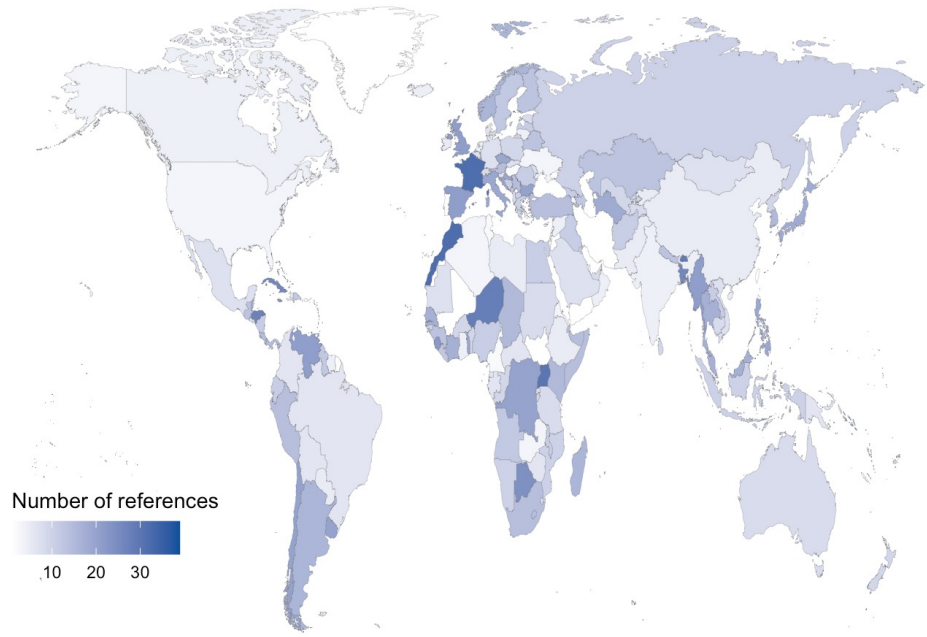
- We want to identify mentions of health implications of climate change.
- One approach: Keyword-in-Context (KWIC)
  1. Identify health references
  2. Look around for climate change references
  3. Count
- In more detail:
  1. Build a dictionary of health references (talk to health scientists)
  2. Build a dictionary of climate change references (talk to climate scientists)
  3. Perform search of health references
  4. Define a window around health references where you look for climate change references +/-5, 10, 15, 20, 25 words.
- A proxy measure that utilises the structure of a formal speech (document) as the foundational assumption.



# Dictionaries

Health terms	Climate change terms
•malaria	•climate change
•diarrhoea	•global warming
•infection	•green house
•disease	•temperature
•sars	•extreme weather
•measles	•global environmental change
•pneumonia	•climate variability
•epidemic	•greenhouse
•pandemic	•low carbon
•public health	•ghge
•health care	•renewable energy
•epidemiology	•carbon emission
•healthcare	•co2 emission
•health	•climate pollutant
•mortality	
•morbidity	
•nutrition	
•illness	
•infectious	
•ncd	
•non-communicable disease	
•noncommunicable disease	
•communicable disease	
•air pollution	
•nutrition	
•malnutrition	
•mental disorder	
•stunting	





# BERT experiments

	precision	recall	f1-score	support
0	0.9826	0.9826	0.9826	4367
1	0.9801	0.9801	0.9801	3828
accuracy			0.9815	8195
macro avg	0.9814	0.9814	0.9814	8195
weighted avg	0.9815	0.9815	0.9815	8195

Total kappa: 0.963

Total MCC: 0.963

## Classification task:

- Train on Wikipedia and WHO (or IPCC) reports
- Finetune on the UNGD corpus
- Predict on sentence level



# Measure validation:

**Togo 2017** (false negative – not picked up in our analysis but clearly about health outcomes of climate change... maybe simple solution is adding term ‘loss of life’ to search terms):

“My country, which often faces the vagaries of coastal erosion, flooding and irregular rainy seasons, believes that the Paris Agreement on Climate Change remains an ideal framework for carrying out joint initiatives that will safeguard the climate.

In West Africa, the recent bad weather, which resulted in significant loss of life in Sierra Leone and tens of thousands of victims in the Niger, has conjured up in people’s memories the spectre of the terrible Ebola crisis that the subregion had to deal with two years ago. Those losses, which we all lament, should compel our individual States and those outside the African continent to act urgently to develop a rigorous and effective plan for ensuring that our peoples have decent living standards and a resilient health-care system, with particular emphasis on our technical services’ early-warning and prevention capabilities.”

# Measure validation:

**Guatemala 2017** (false positive – the reference to climate change is in a separate paragraph but within the “window” to health)

“Guatemala already has a national action plan on climate change, so that our domestic legislation may be carried out to achieve mitigation, adaptation and resilience to climate change.

With regard to nutrition and food security, with the support of the World Food Programme we have put together the final report of the strategic review of the situation in Guatemala in that regard, with an emphasis on chronic malnutrition and its root causes.”

# Measure validation:

**Trinidad and Tobago 2012** (false positive – again the reference to climate change is in a separate paragraph – but next sentence – to the health reference)

That would guarantee the survival of those of us who are most vulnerable to climate change, climate variability and sea-level rise. Every moment of delay takes us closer to our own demise.

The health and well-being of our people is central to our development. The prevalence of non-communicable diseases, in particular, poses a serious threat to Caribbean development.



# Measure validation:

## **Australia 2019** (potential false positive)

“Our goal is simple — to ensure that the south-west Pacific is secure strategically, stable economically, sovereign politically and sustainable environmentally.

The Organization’s work, in partnership with Australia, has also helped to build a more sustainable and resilient Pacific, to support local climate change actions and resilience, to strive for gender equality through the empowerment of women and girls, to support continuing improvements in health outcomes and to bolster regional peace, including through the Bougainville Referendum Support Project.”

- This isn’t a straightforward false positive in that one could argue that broadly the Australian government is linking its work in the Pacific on climate change and improving health, but it’s not a straightforward example of health outcomes of climate change.

# Questions

- Answers to a wrong question?
- We are interested in attribution – a question of causal relationships.

# Lancet Countdown in Europe



- Digital public health surveillance system
- Attributional characterisation of health and climate change link
  - Citizen level (self-reported on social media)
  - Political level (statements in the European and national parliaments)
  - Corporate level (Company CSR statements)
  - Media reports

# Where we are

Annotating real world political speeches along:

1. causal relations as introduced intra-clausally (within the same clause) and
  2. relations that are introduced inter-clausally as a relation between different units of discourse (e.g. sentences)
- We follow the definition of causal language based on BECauSe 1.0 defined as “any construction which presents one event, state, action, or entity as promoting or hindering another, and which includes at least one lexical trigger.”

5K sentences from UNGD corpus (2020 session).



[jankin@hertie-school.org](mailto:jankin@hertie-school.org)  
[@smych](#)