

This is not an ADB material. The views expressed in this document are the views of the author/s and/or their organizations and do not necessarily reflect the views or policies of the Asian Development Bank, or its Board of Governors, or the governments they represent. ADB does not guarantee the accuracy and/or completeness of the material's contents, and accepts no responsibility for any direct or indirect consequence of their use or reliance, whether wholly or partially. Please feel free to contact the authors directly should you have queries.

Asian Impact Webinar

Leveraging Social Media to Discern the Public Voice for Development

Mapping the Public Voice for Development:

Natural Language Processing of Social Media Text Data

A Special Supplement of the Key Indicators for Asia and the Pacific 2022

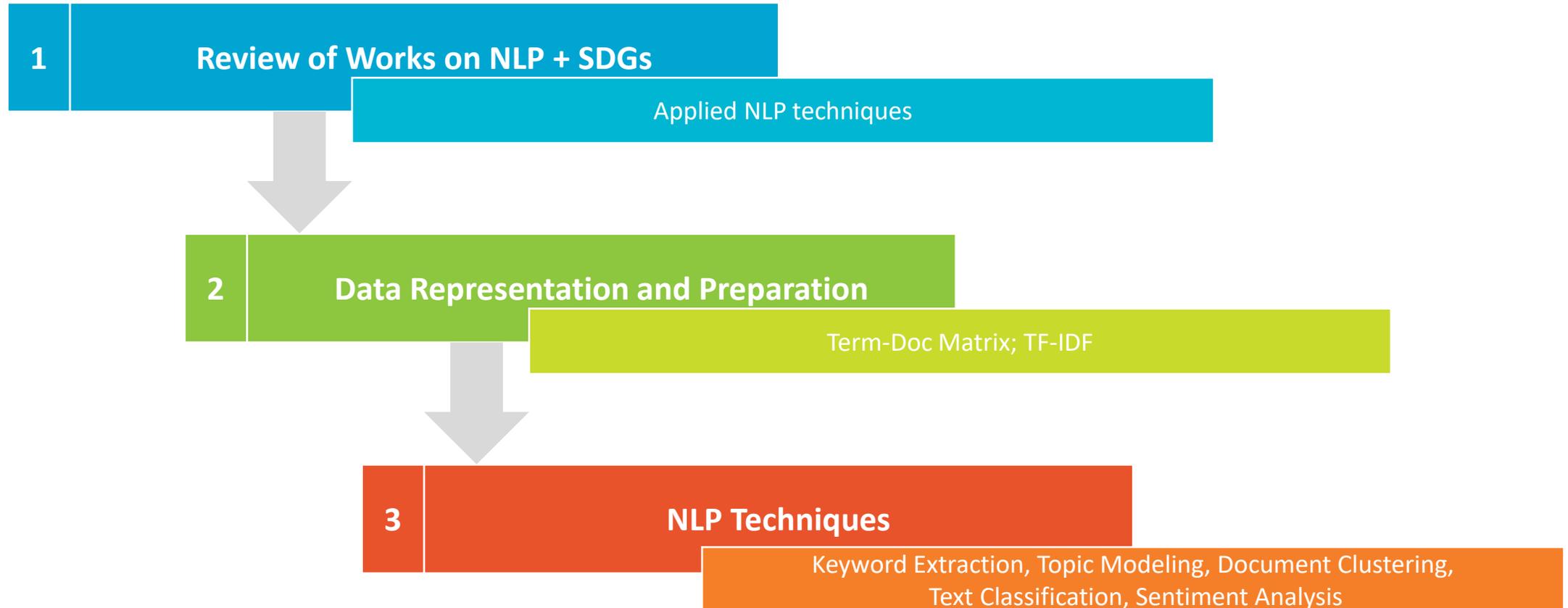
By Stanley Lawrence Sie, Charibeth Cheng, Cedric Basuel, Alyssa Villanueva, Kadra J. Saeed, and Daniel Boller

October 07, 2022

Motivation

- Sound economic research and policy making depend on timely and granular data—the coronavirus disease (COVID-19) pandemic further highlighted the need for real-time and comprehensive information.
- With the advent of social media, allowing individuals to share and receive information, social media text data become a valuable source for timely and granular information.
- Social media text data can be leveraged to map and track the behavior, opinions, concerns, and expectations of citizens, respectively can be used in policymaking to “enhance the quality of government” by improving civic engagement, citizen participation in government, and customer service (Wahyunengseh, Hastjarjo, and Suharto 2018).
- Natural Language Processing (NLP), the computational preparation and analysis of text data, can be applied to study, for example, topics and sentiment in natural language, including social media text data.

Framework of the Report

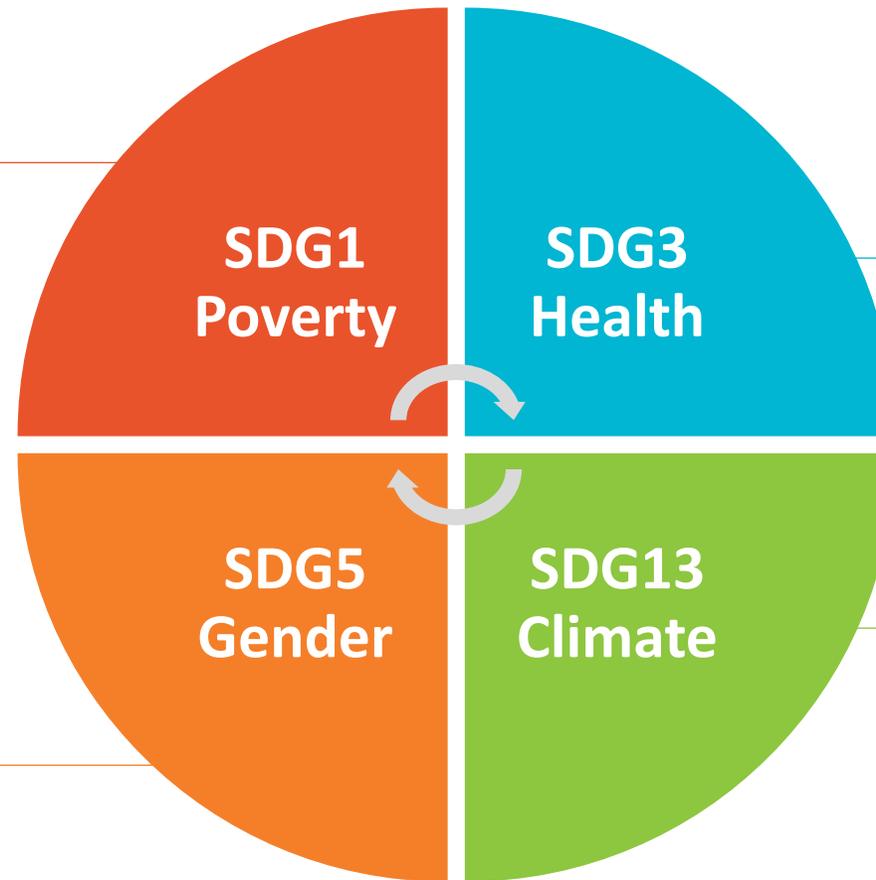


Review of Works on NLP + SDGs

- Identify life satisfaction and worries (family health and employment)

Text Classification and Topic Modeling

Source: Rohrer et al. 2017



- Identify at-risk locations for HIV contamination
- Discover barriers in opioid use disorder treatment
- Determine health literacy among diabetes patients

Text classification, Keyword Extraction, Topic Modeling

Source: Weibel et al. 2017; Xie et al. 2021; Schillinger et al. 2020

- Gather evidence of gender prejudice in languages
- Determine perceptions in the practice of child marriages
- Identify locations where harassments usually occur

Text Classification, Document Clustering, Keyword Extraction, and Sentiment Analysis

Source: Orgeira-Crespo et al. 2021; Syahirah et al. 2018; Alonso-Parra et al. 2021

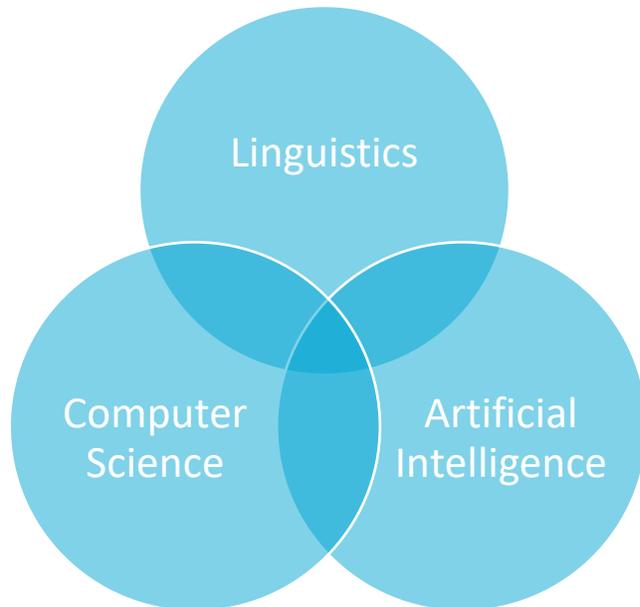
- Identify gaps in climate change awareness among different societal groups
- Gather evidence on hazard risks

Topic Modeling

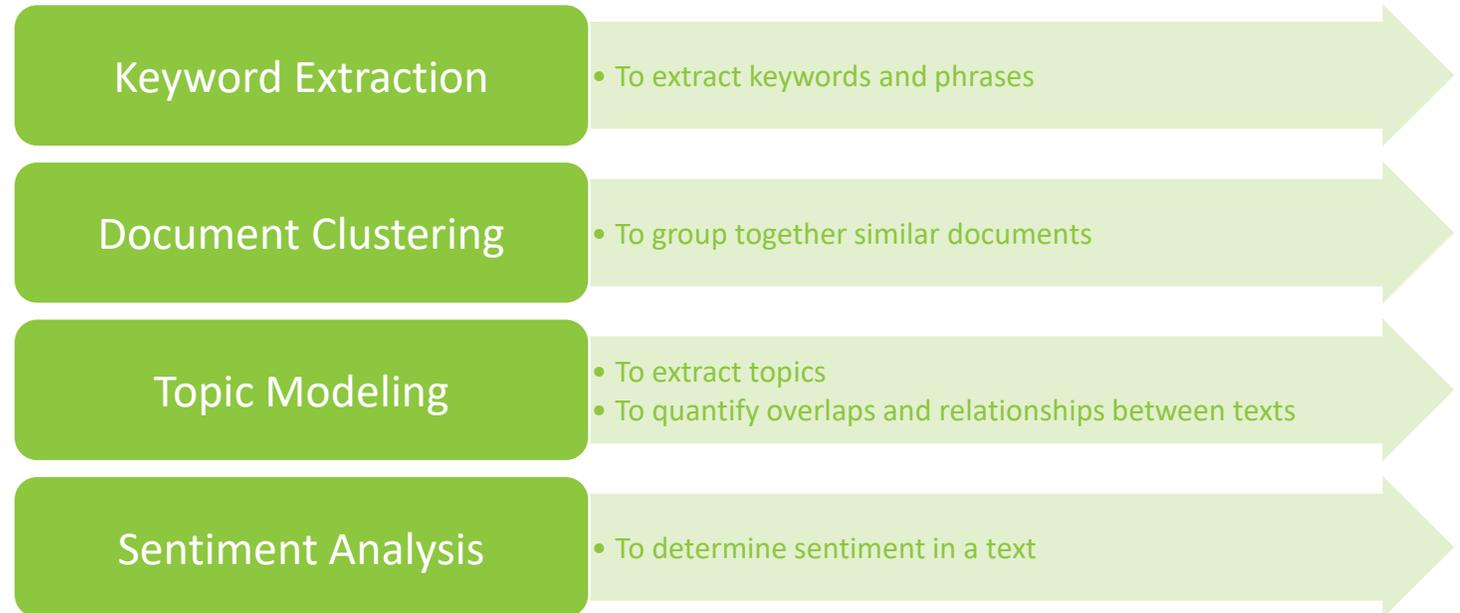
Source: Hwang et al. 2021

Natural Language Processing (NLP) Techniques

NLP dimension

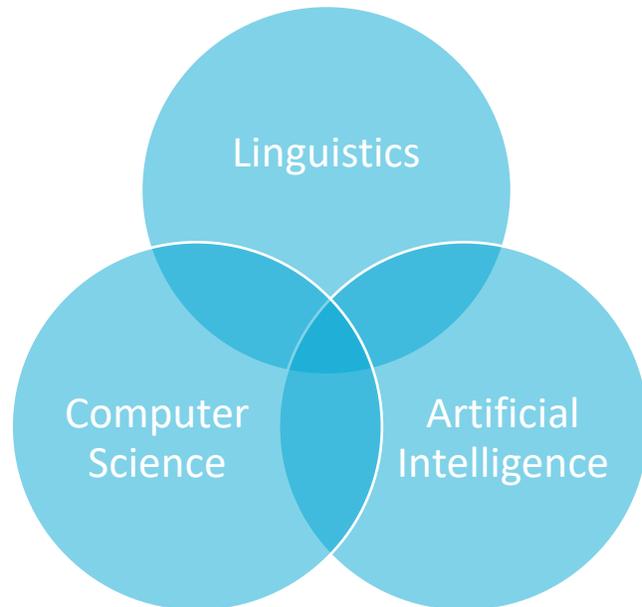


NLP techniques introduced in the report

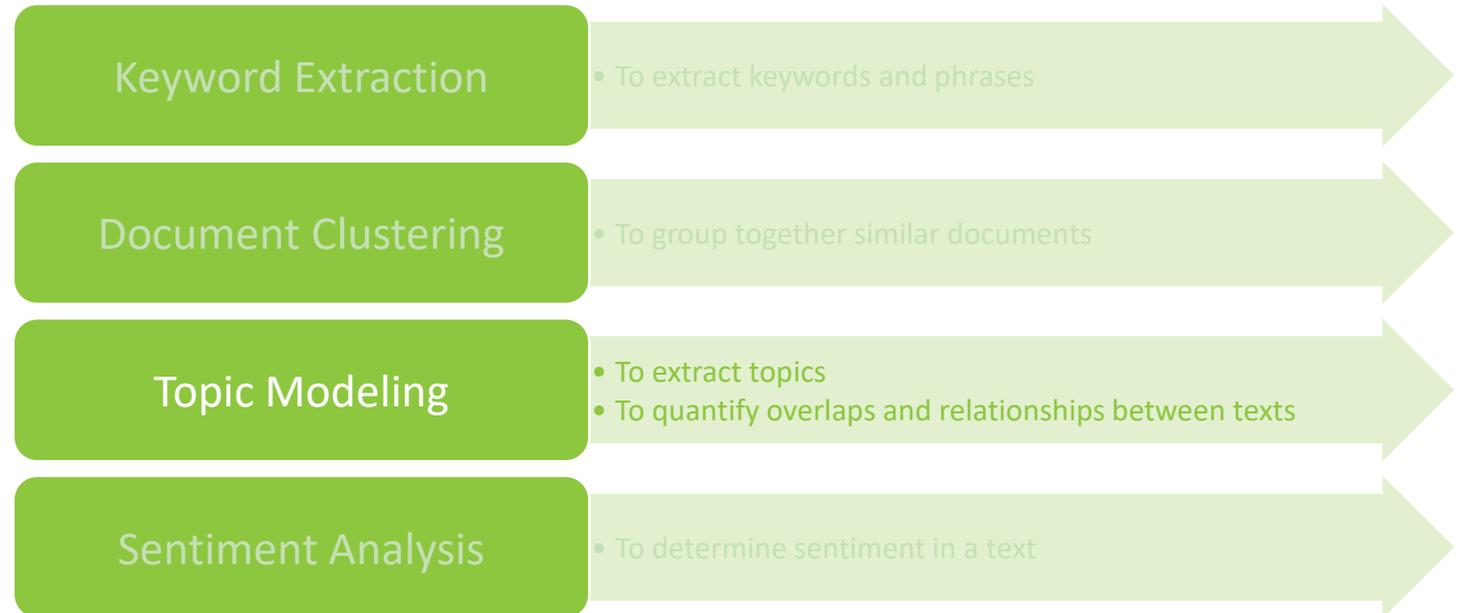


Natural Language Processing (NLP) Techniques

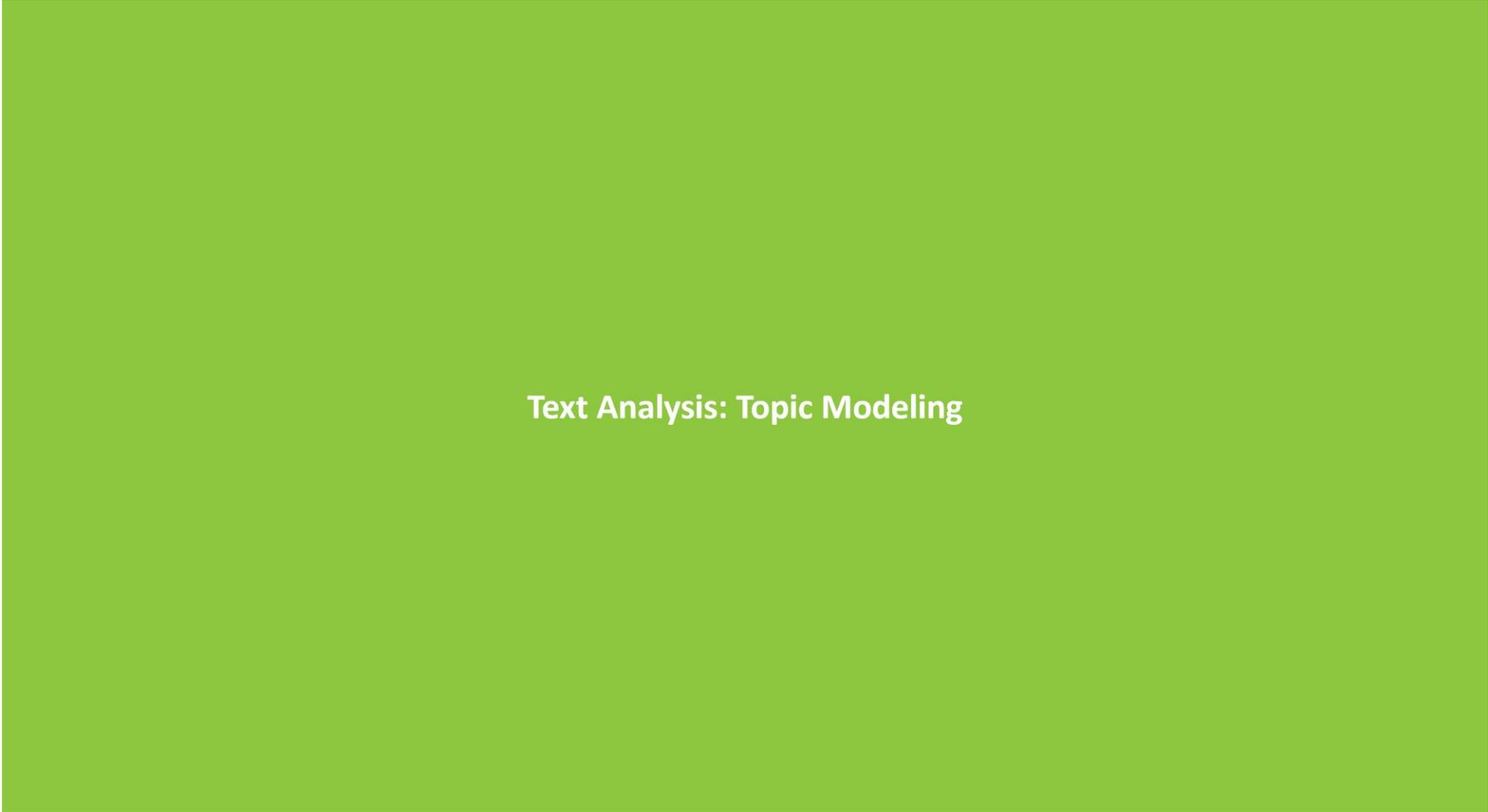
NLP dimension



NLP techniques introduced in the report



Text Analysis: Topic Modeling



Case Study on Climate Change in Australia: Motivation and Data

- Data source: **Twitter Historical Powertrack API**
- Data collection
 - Country: Australia
 - Time: January 15 to December 13, 2021
 - Keywords: climate change, global warming, #climatechange, #globalwarming
- Data Preparation:
 1. Removal of duplicated tweets.
 2. Conversion of non-word tokens (e.g., URLs) into special tokens.
 3. Removal of entire email addresses and special characters (e.g., non-alphanumeric characters).
 4. Performed lemmatization (transforming a text to its root form), spellchecking, and lowercasing.
 5. Identification and removal of stop words to remove common words that appear multiple times within the text corpus.
- Final sample:
 - Total Twitter tweets: 8,302
 - Unique Twitter users: 2,495

Case Study on Climate Change in Australia: Results (I)

Topic Labels

(based on language proficiency and discussions of the research team)

1. Timely action, immediate action, call to action
2. Identify / report needs / problems to government
communication between people & government, find solutions
3. Make science & knowledge actionable
4. Check impact of actions, Australian fires
5. Cost and effect of fossil fuel industry
6. Impact of elections and demand for policy measures /
intervention
7. Temperature changes, threats
8. Responsibility of the leaders, propaganda, climate change
deniers

Topic	Keywords
1	need, action, time, real, great, world, let, big, help, work, happen, bad, covid, way, stop, thing, long, look, lead, lot, save, australian, start, nuclear, thank, pandemic, deal, power, energy, bushfire
2	people, government, right, report, use, problem, come, find, address, issue, need, anti, feel, try, live, hope, leave, vaccine, lack, today, remember, planet, business, care, hot, child, kid, read, state, look
3	know, good, world, science, country, far, continue, scientist, protest, reason, life, thing, government, decision, question, watch, care, mass, carbon, medium, rest, extinction, want, policy, public, debate, turn, nation, nature, charge
4	year, think, impact, fire, love, ago, political, mean, word, research, wonder, bit, number, hold, ignore, contribute, support, past, attend, country, pretty, enjoy, warm, pm, drought, failure, issue, risk, ignorant, mention
5	fuel, fossil, cost, need, effect, coal, disaster, high, govt, event, consequence, industry, gas, huge, world, emission, sea, power, stop, large, true, majority, agree, face, avoid, weather, new, war, reduce, animal
6	vote, plan, want, target, net, time, decade, new, tax, win, emission, increase, policy, late, denial, leadership, expect, like, believe, massive, area, result, coalition, fail, government, achieve, break, meet, stop, pay
7	flood, believe, temperature, sure, theory, level, renewable, set, cold, course, early, mean, wake, politician, work, threaten, control, planet, discuss, society, rise, low, scientific, cool, apparently, corporate, export, today, threat, group
8	leader, fact, health, actually, lie, talk, politic, denier, technology, election, drive, fool, lose, little, elect, view, labor, [Removed Word], boy, poor, policy, trust, absolutely, list, read, pm, carbon, propaganda, management, pandemic

Case Study on Climate Change in Australia: Results (II)

Topic Labels

(based on language proficiency and discussions of the research team)

- 9. Protecting threatened jobs, job security
- 10. Solution-finding for biodiversity / environment loss & damage
- 11. Economic factors, spending
- 12. Women's advocacy for climate change
- 13. Human concerns / fears / worries on climate change
- 14. Action vs inaction on climate change
- 15. Truth-seeking, debate about the truth on climate change

Topic	Keywords
9	future, job, hear, threat, protect, away, support, grow, sign, security, build, friend, invest, join, voter, stay, land, food, family, want, deni, action, fight, guy, piece, tech, safe, worker, rock, rate
10	run, understand, party, crisis, follow, environment, damage, water, environmental, existential, amazing, biodiversity, responsible, problem, reduce, solve, loss, happy, paper, concern, modelling, solution, generation, life, driver, policy, single, issue, kind, conversation
11	money, cause, news, destroy, economy, ask, fund, matter, conservative, rich, story, allow, fall, dead, mind, spend, die, red, public, study, refuse, elite, self, respond, rule, reminder, left, end, accord, deserve
12	deny, woman, pollution, include, speak, week, catastrophic, ecosystem, learn, emergency, [Removed Word], appear, possible, wonderful, advocate, day, attitude, plastic, year, drop, farm, wear, river, health, imagine, publish, nd, funding, value, product
13	human, record, worry, approach, car, fear, example, thread, form, worried, activity, support, healthcare, sustainable, undermine, induce, fine, type, test, urgent, ok, acknowledge, wealth, fix, seek, year, ridiculous, beat, radical, petrol
14	act, inaction, earth, maybe, summer, natural, role, answer, warn, floor, population, totally, catastrophe, embarrassing, cross, honest, useless, office, code, good, agenda, probably, line, recent, improve, arrive, demonstrate, spread, actively, pro
15	truth, position, million, measure, share, game, room, factor, apply, prediction, argument, [Removed Word], look, sort, religion, brilliant, influence, shoot, bias, researcher, urban, design, specie, snow, elephant, tackle, item, doom, sociopath, death

Case Study on Climate Change in Australia: Implications / Discussion

- **LDA topic models permit a comprehensive mapping of topics**
 - Identification of patterns (topics) in unstructured (social media) text data
 - Enables drawing of conclusion about communication in social media
- Public discourse on climate change in Australia **focuses on policy making and contributions of society**
 - Topic 2 (identify / report needs/problems to government, communication between people and government / find solutions)
 - Topic 6 (impact of elections and demand for policy measures / intervention)
 - Topic 15 (truth-seeking / debate about the truth on climate change)
- Identified topics on climate change for Australia allow not only to **reflect on concerns, ideas, and expectations of the public**, but also to **develop approaches and measures to tackle climate change**.
- Development of approaches and measures to tackle climate change **necessitates a joint dialogue between all social groups** and **remains a fundamental responsibility of the entire society**.

Asian Impact Webinar

Leveraging Social Media to Discern the Public Voice for Development

Mapping the Public Voice for Development:

Natural Language Processing of Social Media Text Data

A Special Supplement of Key Indicators for Asia and the Pacific 2022

By Stanley Lawrence Sie, Charibeth Cheng, Cedric Basuel, Alyssa Villanueva, Kadra J. Saeed, and Daniel Boller

October 07, 2022