Climate Conversations: A Policymaker's Toolkit

Kate Habich, Kathryn Link-Oberstar,
Shay Milner, Jennifer Yeaton

Significance of Climate Communications

- *Public sentiment* plays an important role shaping responses to climate change
- Climate advocates are tasked with communicating *factual* and *compelling* information
- Reactions to climate information is shaped by cultural background, personal experiences, ingrained beliefs, attitudes, and values
- Tailoring climate communications to resonate with these deeply held values is essential for ensuring the message is effectively received by the intended audience.

Project Overview

We created a **policymaker's toolkit** to help policy makers effectively communicate with their constituents about climate change, focusing on:

1. Coverage of topics their constituents are most interested in
2. Accurate delivery of scientific information
3. Effective communication in line with values constituents ascribe to

Project Goals...

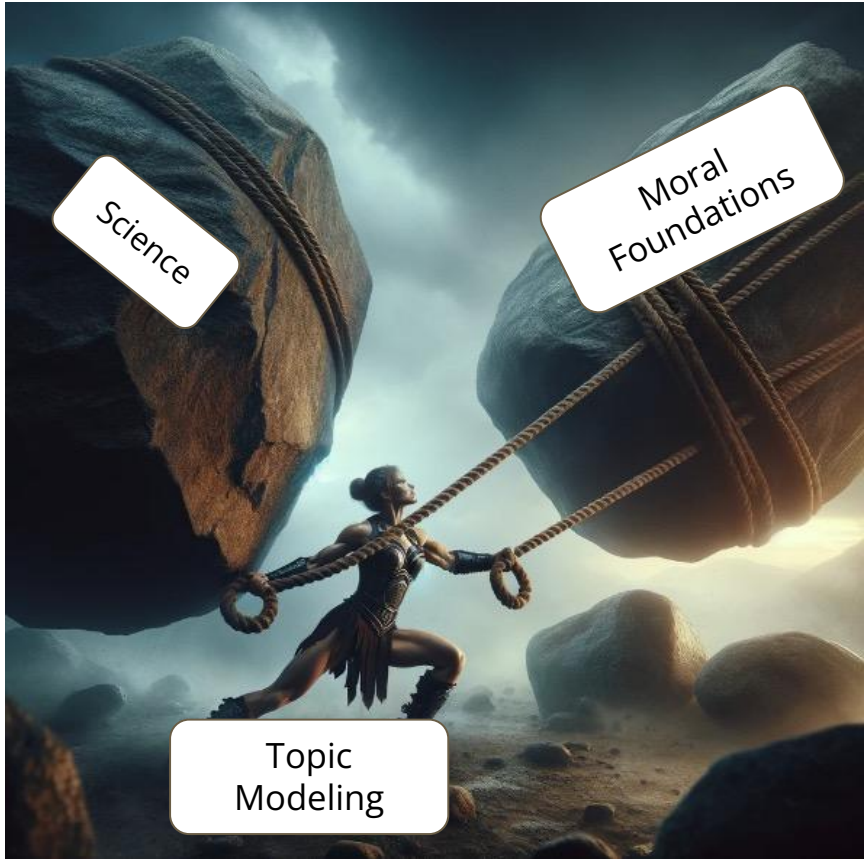


Image courtesy of DALL-E

Sections

- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

Sections

- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

Scientific Article Data Collection

- Used ~10,000 articles in the fields of environmental science and environmental policy to train article summarizers
- Pulled from ProQuest using TDMStudios tool
- Articles chosen from journals with impact scores >1 to emphasize impactful and well-written journals for training
- All articles collected were published after Jan 1, 2023 to maintain scientific relevance

Reddit Data Collection

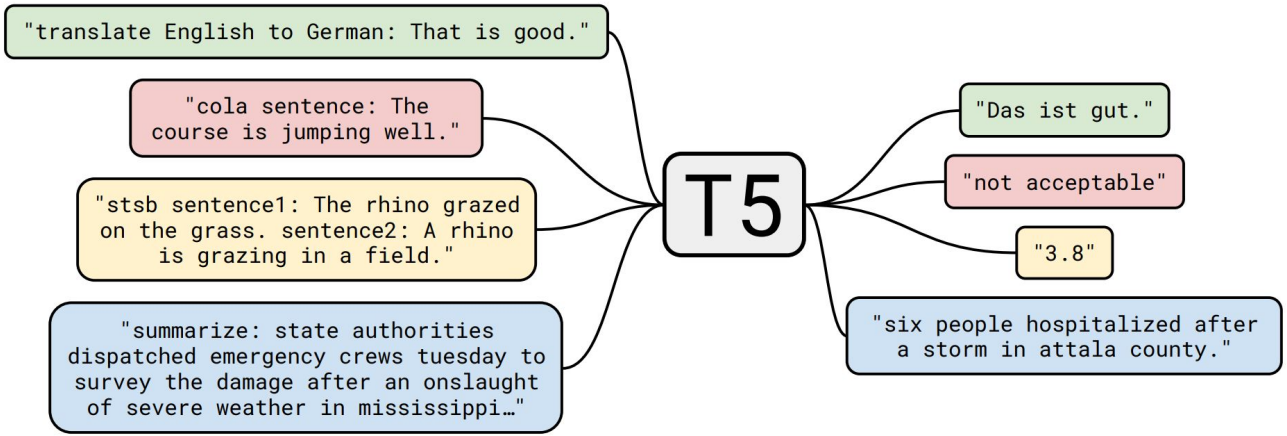
- Focused on subreddits focused on climate change:
  - r/climate
  - r/environment
  - r/climatechange
  - r/climateskeptics
  - r/climateOffensive
- Along with broader subreddits where discussion of climate change occurs:
  - r/science
  - r/politics
  - r/worldnews
- Pulled data from the 2 weeks prior, week during, and 2 weeks after a major climate event (Hurricane Ian in 2022)
  - Pre: September 9 - 22, 2022
  - During: September 23 - 30, 2022
  - Post: October 1 - 14, 2022
- Size of dataset: about 4 million comments and 55,000 submissions

Sections

- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

Abstract Summarization Model

- Adapted HuggingFace abstractive summarization model
  - Batch size = 8, epochs = 1, optimizer = AdamWeightDecay
- Trained on the 10,000 scientific article abstracts and titles to produce short, one-line summaries



Abstractive Summarization Model Example and Results

- Model is evaluated based on a Recall-Oriented Understudy for Gisting Evaluation (ROUGE) score

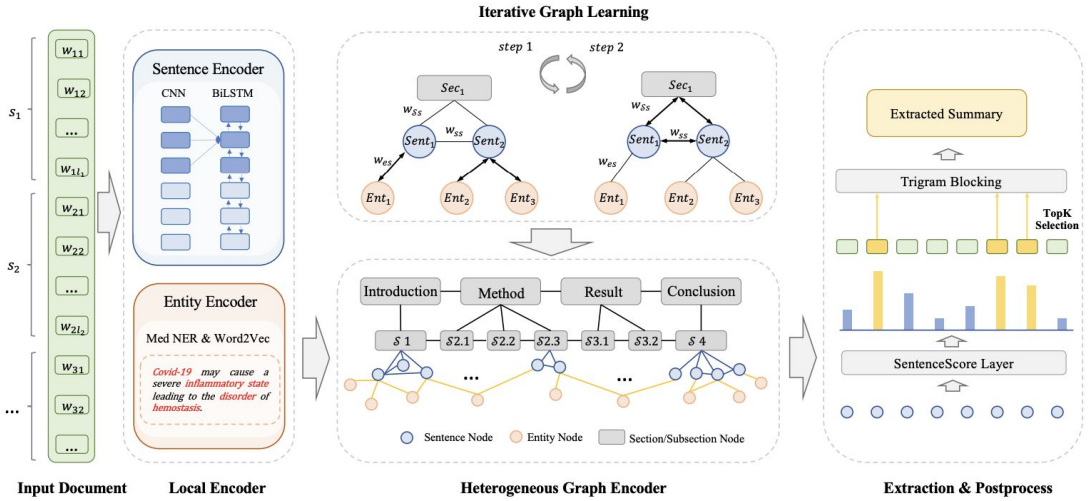
ROUGE Score combines precision and recall to describe the overlap of n-grams in the reference and generated text

  - Unigram ROUGE Score: 38.6996
  - Bigram ROUGE Score: 18.6389
  - ROUGE-L Score: 32.1834
- Model Loss
  - Before fine-tuning: 2.6171
  - After fine-tuning: 2.2893

Extractive Summarization: SAPGraph

Original approach

- Implementation SAPGraph
- Heterogeneous graph architecture
- Challenge of summarizing complex scientific texts
- Improves summarization performance

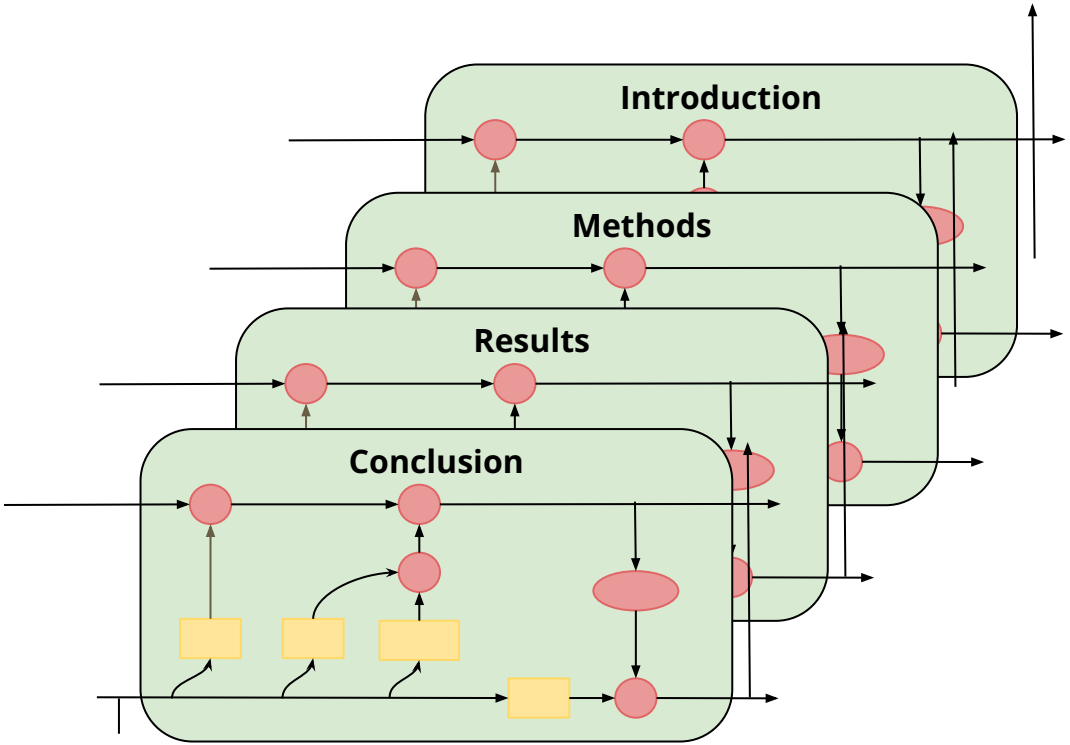


From Qi, Siya et al. 2022. "SAPGraph: Structure-aware Extractive Summarization for Scientific Papers with Heterogeneous Graph." *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 12th International Joint Conference on Natural Language Processing*, vol 1, pp. 575-86.

Extractive Summarization: Method

Revised approach

- Challenge implementing SAPGraph
- Replicate method of isolating article sections
- Developed four Long Short-Term Memory (LSTM) models



Extractive Summarization: Process

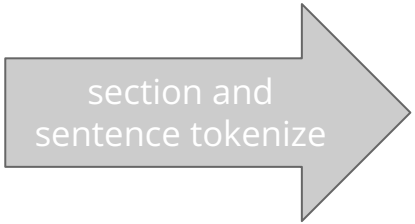
Introduction
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Suspendisse arcu risus, malesuada ac fermentum at, posuere elit.

Methods
Vivamus sagittis molestie elit. Aenean augue dolor, ultricies id justo non, fringilla blandit elit. Nullam id mauris a erat.

Results
Mauris lacus ipsum, sollicitudin vel mi vitae, gravida sodales sem.

Conclusion
Nunc venenatis elit in libero finibus, vitae aliquam mi egestas. Donec interdum quam mattis ante egestas imperdiet.

References
Fusce Fringilla. 2024. Bibendum metus vitae tempor. *Sed sollicitudin erat.*



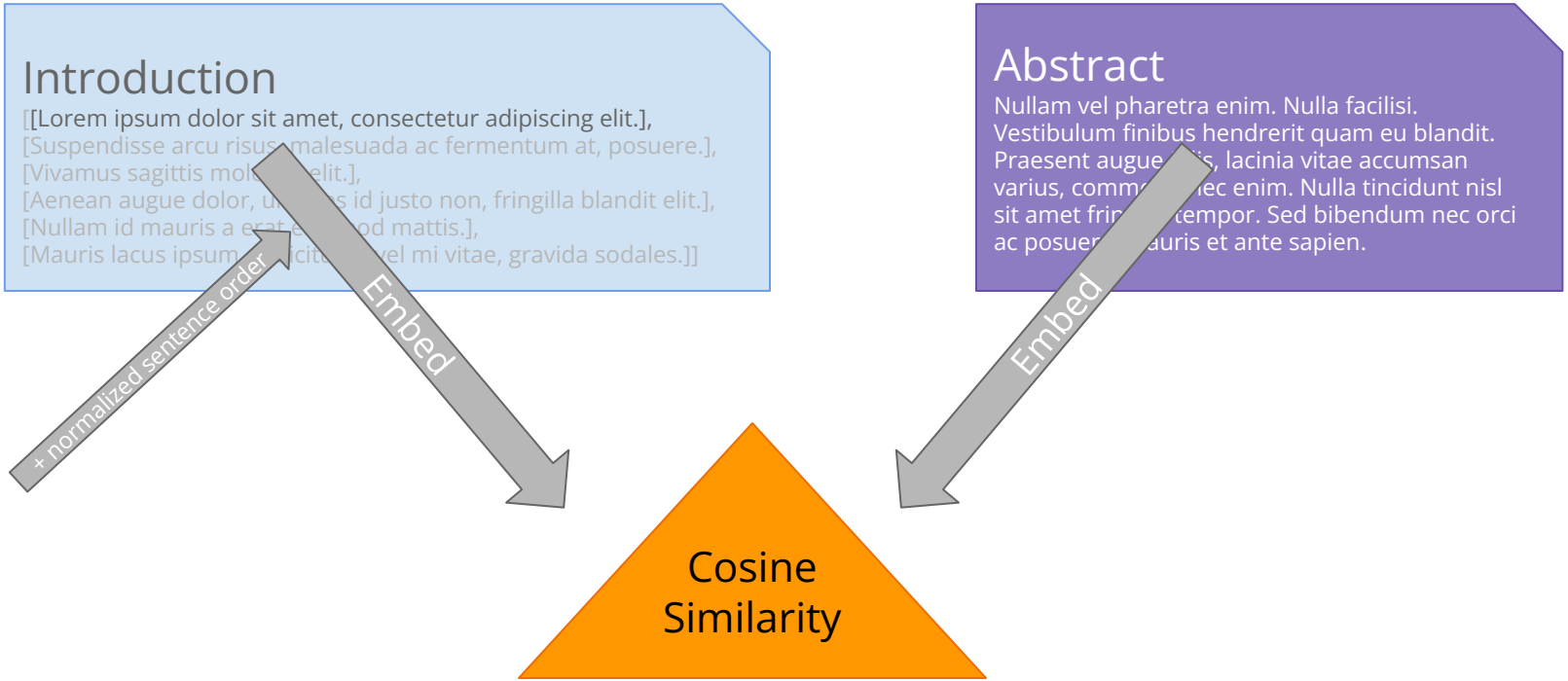
Introduction
[[Lorem ipsum dolor sit amet, consectetur adipiscing elit.], [Suspendisse arcu risus, malesuada ac fermentum at, posuere.], ...]

Methods
[[Vivamus sagittis molestie elit.], [Aenean augue dolor, ultricies id justo non, fringilla blandit elit.], [Nullam id mauris a erat.], ...]

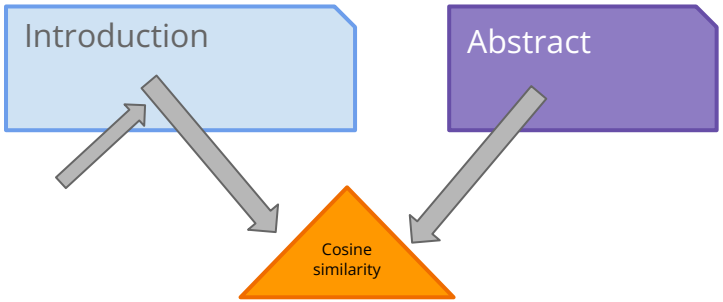
Results
[[Mauris lacus ipsum, sollicitudin vel mi vitae, gravida sodales sem.], [Vestibulum feugiat commodo est sed dapibus.], ...]

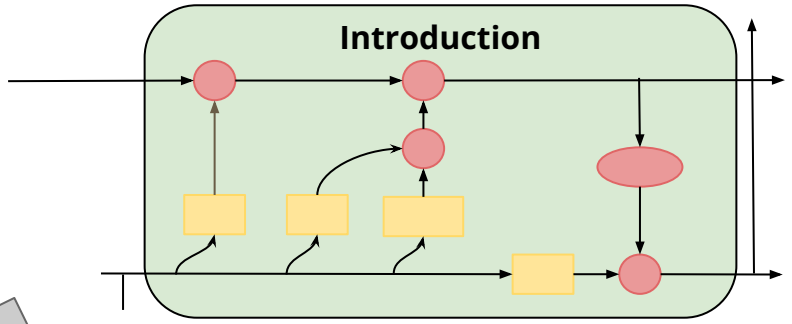
Conclusion
[[Nunc venenatis elit in libero finibus, vitae aliquam mi egestas.], [Donec interdum quam mattis ante egestas imperdiet.,]...]

Extractive Summarization: Model Development



Extractive Summarization: Model Development





Importance Scores

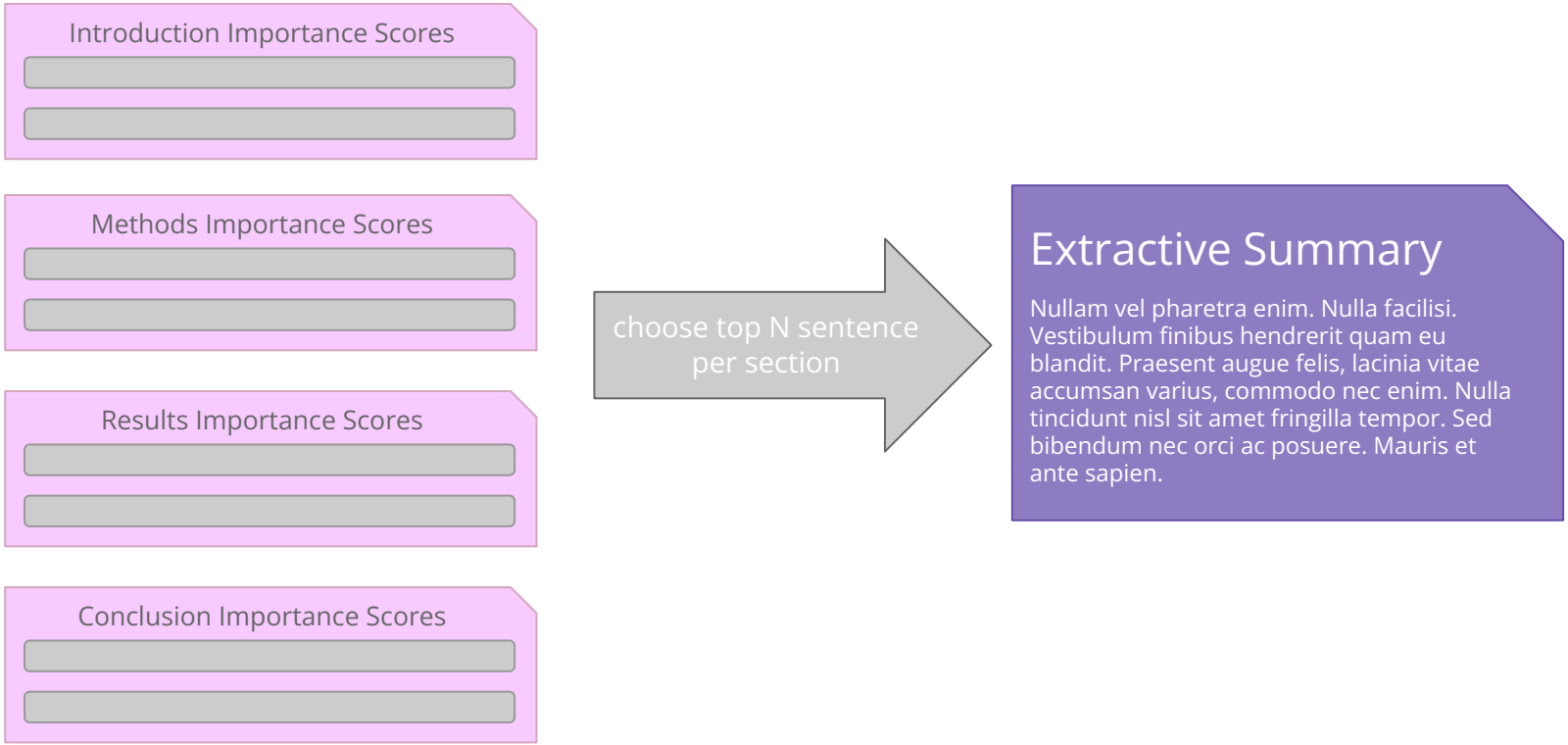
(Lorem ipsum dolor sit amet, consectetur adipiscing elit., 0.67)

(Suspendisse arcu risus, malesuada ac fermentum at., 0.88)

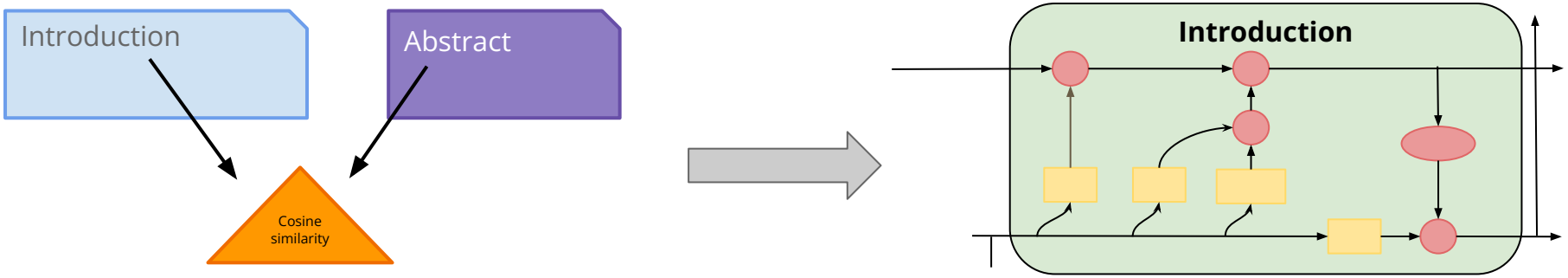
(Vivamus sagittis molestie elit., 0.43)

(Mauris lacus ipsum, sollicitudin vel mi vitae, gravida sodales., 0.29)

Extractive Summarization: Summarization



Extractive Summarization: Model Development



Model	Initial Validation to Final Test Loss
Introduction	0.0331 → 0.0137
Methods	0.0297 → 0.0123
Results	0.0344 → 0.0131
Conclusions	0.0398 → 0.0101

Extractive Summarization: Example

Abstract

Abrupt monsoon onsets/retreats are indispensable targets for climate prediction and future projection, but the origins of their abruptness remain elusive. This study establishes the existence of three climatological Madden-Julian Oscillation (CMJO) episodes contributing to the rapid Australian summer monsoon retreat in mid-March, the South China Sea (or East Asian) summer monsoon onset in mid-May, and the Indian summer monsoon onset in early June. The CMJO displays a dynamically coherent convection-circulation structure resembling its transitionary counterpart, demonstrating its robustness as a convectively coupled circulation system and the tendency of the transient MJOs' phase-lock to the annual cycle. The CMJO is inactive during the boreal winter due to destructive year-to-year modulations of El Niño-Southern Oscillation. We hypothesize that the interaction between atmospheric internal variability (MJO) and the insolation-forced slow annual cycle generates the sudden monsoon withdrawal/onset during the boreal spring. Understanding the factors determining the timing and location of the MJO's phase-locking and its variability is vital for monsoon forecasting and climate projection.

Summarization

Improving the understanding of the sudden changes and monsoon singularities in the annual cycle has profound implications for the seasonal forecasts and climate projections, as socio-economic activities, including agricultural planning and disaster mitigation, demand accurate monsoon onset and withdrawal predictions both presently and in the future. The summertime monsoonal CISO has been widely recognized to shape the monsoon onsets and retreats in East Asia, WNP, and South Asia (e.g., Ref. Daily atmospheric diabatic heating and sea surface temperature (SST) in the WNP might be important factors in maintaining the CISO related to the EA summer monsoon (EASM)16. Predicting monsoon onset, peak, and withdrawal has profound implications on socio-economic activity and natural disaster management51,52. A better understanding of the CISO and its relationship with the monsoon life cycle can advance our knowledge of the source of the S2S predictability and improve weather and climate predictions. Climate models' capability in simulating and predicting CISO and the abrupt monsoon transition remains ample room for improvement55–58. Through examining a 227-year daily precipitation record gathered in Seoul, South Korea, the characteristics of the rainy season, including the dates of onset, retreat, and summit, exhibit significant centennial variations, indicating the time- varying monsoon singularities detected by a 30-yr climatology change with time14. Future studies of the secular changes in CISO and monsoon singularity will help better understand the cause of climate change. Kajikawa, Y. and Yasunari, T. Interannual variability of the 10–25- and 30–60-day variation over the South China Sea during boreal summer. Dai, L., Cheng, T. F. and Lu, M. Anthropogenic warming disrupts intraseasonal monsoon stages and brings dry-get-wetter climate in future East Asia.

Sections

- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

Moral Foundations for Climate Communications

- Moral Foundations provide a mechanism to categorize people's *existing moral frameworks*
- *Five foundations* that comprise the *building blocks of morality* regardless of the culture
- Five fundamental moral foundations:
    1. Care & Harm
    2. Fairness & Reciprocity
    3. Ingroup & Loyalty
    4. Authority & Respect
    5. Purity & Sanctity
- For classification tasks, we divide each foundation into two categories: 'virtue' and 'vice', for 10 total categories

Applications of Moral Foundations

- Moral foundations have been applied across disciplines, particularly powerful in the study of *political ideologies*
- Increasingly in Natural Language Processing as a way to *classify the moral sentiments of different texts*
- **González-Santos et al. (Knowledge Based Systems, 2023):**
  - Automatic Assignment of Moral Foundations to Movie Synopsis
  - Used Word2Vec & BERT to assign moral foundations by comparing tags (tokens) to the Moral Foundations Dictionary as well as an Expanded MFD created using embeddings
  - Demonstrates the applicability to automate moral foundation classification using NLP (formerly done manually)
  - Word2Vec outperformed BERT in their analysis

Evaluating Moral Foundations of Reddit Comments

Part 1: Moral Foundation Dictionary Generator

- Baseline: Moral Foundation Dictionary, crafted by the originators of Moral Foundations Theory with 11 distinct groups
- Expanded Dictionary:
  - Used pretrained fse/glove-twitter-200, trained on 2 billion tweets
    - Expand dictionary based on similarity
    - Tailor the conversation to online discourse
  - Experimented with dictionaries of size: 300, 600, 900, and 2,000 and words
  - Settled on dictionary of 300 words (30 per foundation) based on qualitative review & precedent from González-Santos et al. (2023)

```
%
01                  HarmVirtue
02                  HarmVice
03                  FairnessVirtue
04                  FairnessVice
05                  IngroupVirtue
06                  IngroupVice
07                  AuthorityVirtue
08                  AuthorityVice
09                  PurityVirtue
10                  PurityVice
11                  MoralityGeneral
%

safe*                 01
peace*                01
compassion*           01
empath*               01
sympath*              01
care                  01
caring                01
protect*              01
shield                01
shelter               01
amity                 01
secur*                01
benefit*              01
defen*                01
guard*                01
preserve              01 07 09
```

Moral Foundation Dictionary

Example: Expanded Moral Foundations Dictionary

Authority & Respect	**Vice:** betray, dissent, disrespect, nonconformist, lawless **Virtue:** preserve, law, honor, respect, tradition
Fairness & Reciprocity	**Vice:** unfair, unequal, bias, exclusion, bigot **Virtue:** justice, equity, tolerant, impartial, fair
Care & Harm	**Vice:** suffer, kill, crush, abuse, abandon **Virtue:** peace, compassion, protect, shield, preserve
Ingroup & Loyalty	**Vice:** treason, traitor, deserter, imposter, spy **Virtue:** familial, communal, ally, collective, joint
Purity & Sanctity	**Vice:** destroy, exploit, heretic, unclean, disease **Virtue:** preserve, piety, wholesome, integrity, sacred

Evaluating Moral Foundations of Reddit Comments

Part 2: Evaluating & Classifying Reddit Comments in Climate Subreddits

- Dataset:
  - **Subreddits:** r/climate, r/science, r/climatechange, r/climateoffensive, r/environment, r/worldnews, r/climateskeptics, and r/politic
  - **Time Period:** Given the tendency for online climate discourse to spike during major environmental events, we focused on the period immediately before, during, and after Hurricane Ian (September 23, 2022 – September 30, 2022).
- Evaluation:
  - Compared each word in the comment string to each word in the expanded moral foundation dictionary using pre-trained word vectors from fse/glove-twitter-200
  - Set threshold of 0.25 similarity for considering a word.
  - Each comment was assigned 10 scores, ranging from 0 to 1, for each of the 10 moral foundations based on its similarity to the defining words of each foundation.
  - Aggregate to 5 main categories and assign each comment a **Dominant Moral Foundation**

Example: Processed Comments

Subreddit
r/Climate

Comment:
'How much cleaner or not cleaner is natural gas than coal?'

Word Tokens
[clean, clean, natural, gas, coal]

Dominant Moral Foundation
Purity

Subreddit
r/climateoffensive

Comment:
'Don't listen to the haters on this thread. You're on the right track! The quicker we can get airplanes out of our culture, the better.'

Word Tokens
['listen', 'hater', 'thread', 'right', 'track', 'quick', 'airplane', 'culture', 'well']

Dominant Moral Foundation
Care & Harm

Example: Processed Comments

Subreddit
r/Climate

Comment:
'How much cleaner or not cleaner is natural gas than coal?'

Word Tokens
[clean, clean, natural, gas, coal]

Dominant Moral Foundation
Purity

Aggregated Score*

Care	0.346513
Authority	0.342758
Purity	0.369309
Fairness	0.310201
Ingroup	0.318247

** Average cosine similarity*

Full Score*

AuthorityVice	0.310805
AuthorityVirtue	0.314101
FairnessVice	0.280141
FairnessVirtue	0.34026
HarmVice	0.318314
HarmVirtue	0.374712
IngroupVice	0.313413
IngroupVirtue	0.323082
PurityVice	0.345819
PurityVirtue	0.392799

Example: Processed Comments

Subreddit

r/climateoffensive

Comment:

'Don't listen to the haters on this thread. You're on the right track! The quicker we can get airplanes out of our culture, the better.'

Word Tokens

['listen', 'hater', 'thread', 'right', 'track', 'quick', 'airplane', 'culture', 'well']

Dominant Moral Foundation

Care & Harm

Aggregated Score*

Care	0.346513
Authority	0.342758
Purity	0.369309
Fairness	0.310201
Ingroup	0.318247

** Average cosine similarity*

Full Score*

AuthorityVice	0.310805
AuthorityVirtue	0.314101
FairnessVice	0.280141
FairnessVirtue	0.34026
HarmVice	0.318314
HarmVirtue	0.374712
IngroupVice	0.313413
IngroupVirtue	0.323082
PurityVice	0.345819
PurityVirtue	0.392799

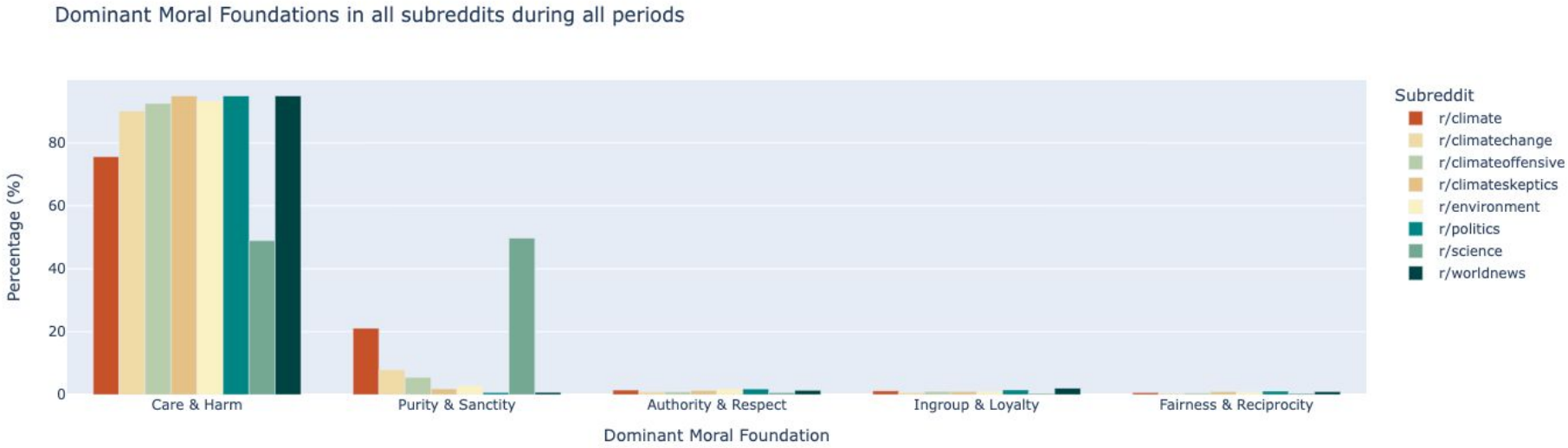
Results

1. Discussions related to **harm and care** are prevalent across all analyzed climate-related subreddits.

2. The language of **purity and sanctity** was **particularly notable in climate and science subreddits**

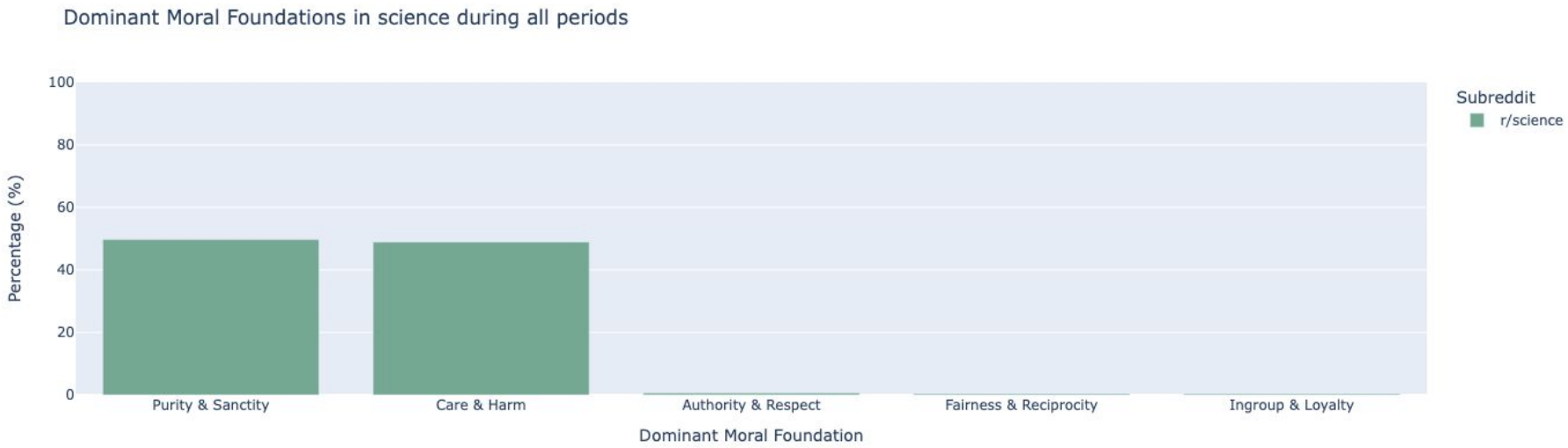
3. **During Hurricane Ian, we observed a spike in the usage of Purity & Sanctity-related terms** across nearly all subreddits, which later reverted to pre-disaster levels

Finding 1: Discussions related to harm and care* are prevalent across all analyzed climate-related subreddits.

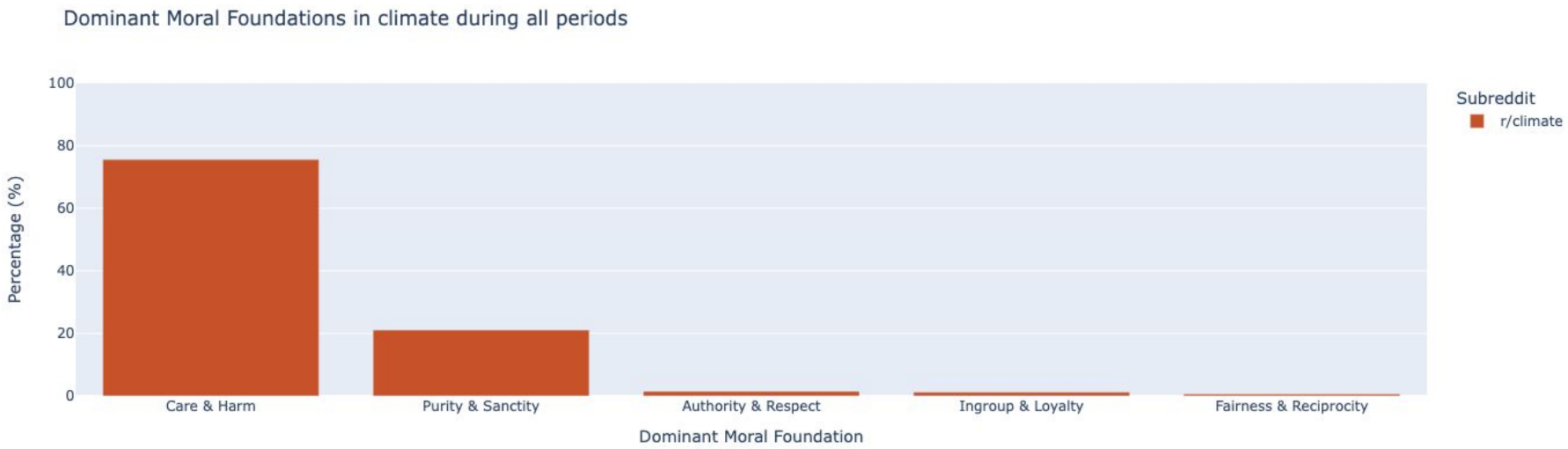


* This includes words like protect, shield, shelter, guard, preserve, suffer, kill, and endanger.

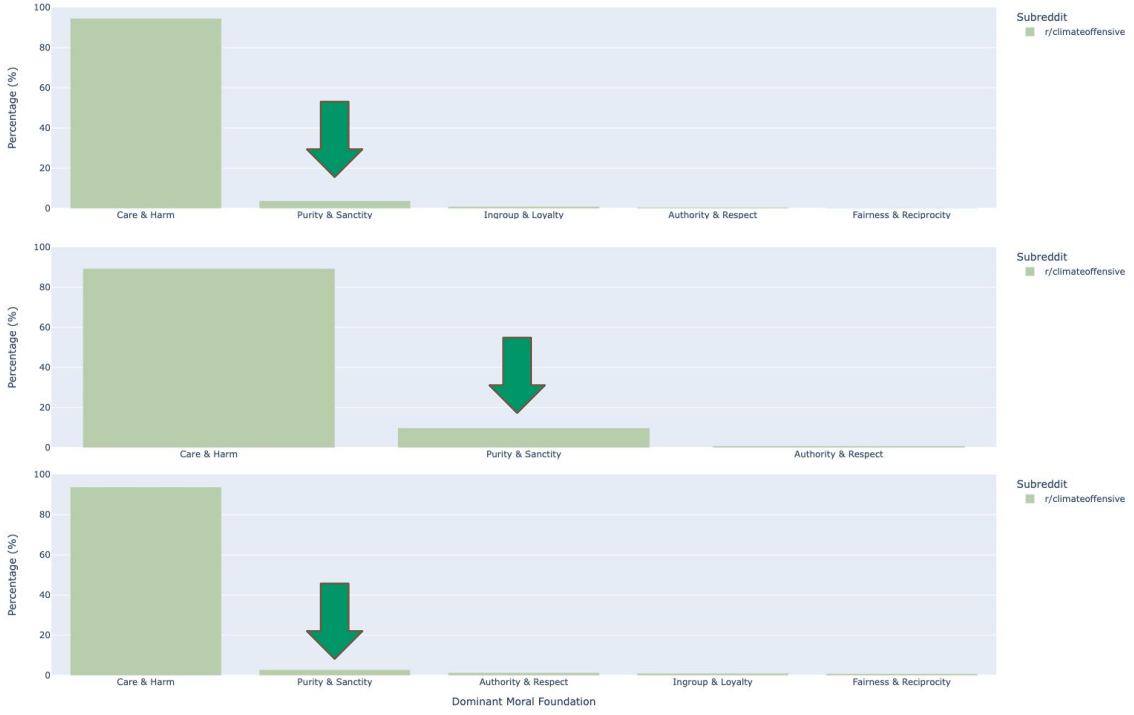
Finding 2: Language of purity and sanctity was particularly notable in climate and science subreddits



Finding 2: Language of purity and sanctity was particularly notable in climate and science subreddits



Finding 3: During Hurricane Ian, spike in the usage of Purity & Sanctity-related terms across nearly all subreddits, which later reverted to pre-disaster levels



Before Hurricane
Purity & Sanctity: 3.8%

During Hurricane
Purity & Sanctity: 9.8%

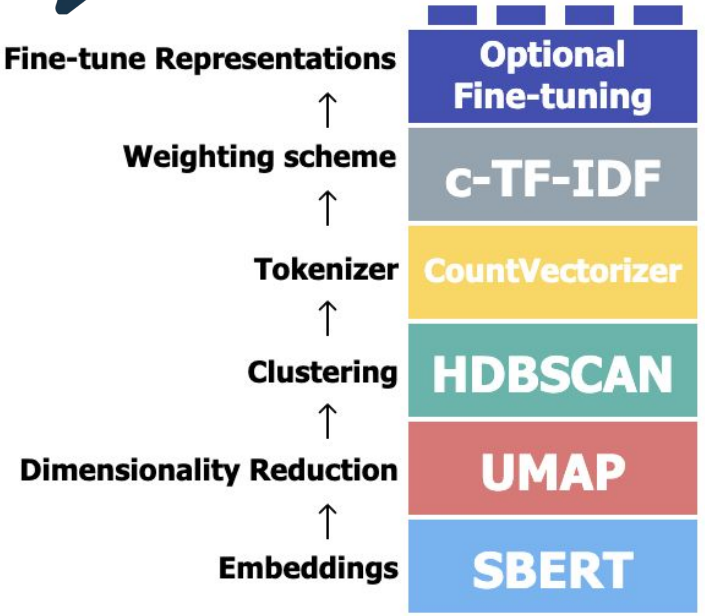
After Hurricane
Purity & Sanctity: 2.9%

Sections

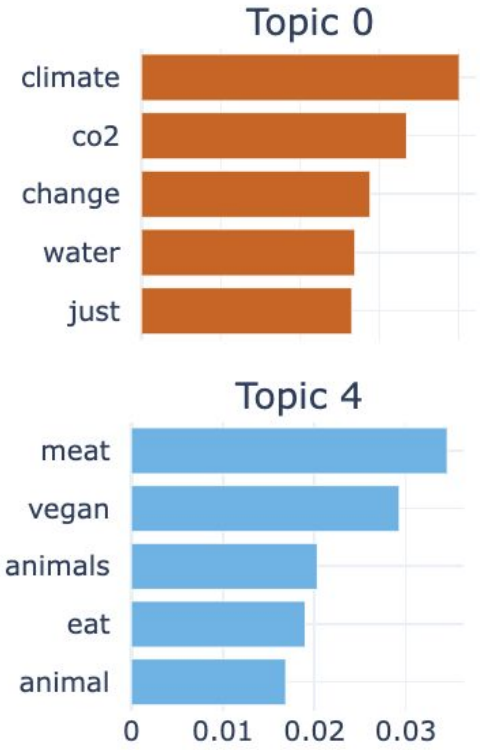
- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

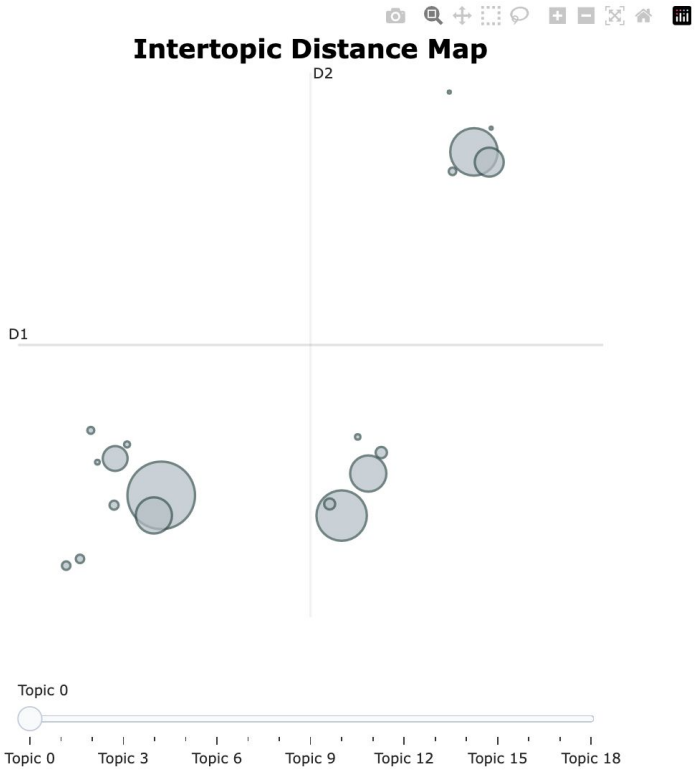
Reddit Topic Modeling: LDA and BERTopic

- LDA:
  - Tokenization
  - Stop word removal
  - Stemming and lemmatization
  - Creating bigrams and trigrams
  - Training LDA model
- BERTopic
  - Pre-processing not required
  - Performs better with shorter text
  - Keywords more interpretable



Reddit Topic Modeling: BERTopic Findings





Reddit Topic Modeling: Ties to other parts of the project

- Collection of words representing our top 10 topics given to moral foundations classifier to determine which moral foundations best match onto each topic
- Topics used to informed which journal articles we pull

Sections

- Data Collection
- Scientific Article Summarization
- Reddit Moral Foundations
- Reddit Topic Modeling
- Toolkit Generation

Toolkit Generation: Bringing it all together

For each topic of interest, we

(1) select the two most pertinent moral foundations to that topic

(2) match an embedding of that topic to keywords in scientific articles

(3) generate short and long summaries of journal articles most pertinent to each topic

(4) compile moral foundations and journal article summaries into a toolkit policymakers can use to communicate accurate information to their constituents about the topics that interest them

Example Output:

```
The representative words for this topic area are ['people', 'just', 'climate', 'like', 'dont', 'change', 'think', 'im', 'years', 'thats']
.The top two moral foundations to focus on when communicating about this topic area are HarmVirtue and FairnessVirtue

.The top 3 journal articles most representative of this topic area with their short and long summaries are:

Short summary:
 <pad> Carbon ecological security framework for regional green and low-carbon development</s>
 .Short summary:
 <pad> Impact of Low-carbon city pilot policy on FDI inflows: a multi-period model and intermediary model</s>
 .Short summary:
 <pad> Impact of Low-carbon city pilot policy on FDI inflows: a multi-period model and intermediary model</s>
 .Short summary:
 <pad> Carbon ecological security framework for regional green and low-carbon development</s>
 .
```

Dashboard of Reddit Findings

Climate Conversations: Dominant Moral Foundations Across Climate Related Subreddits

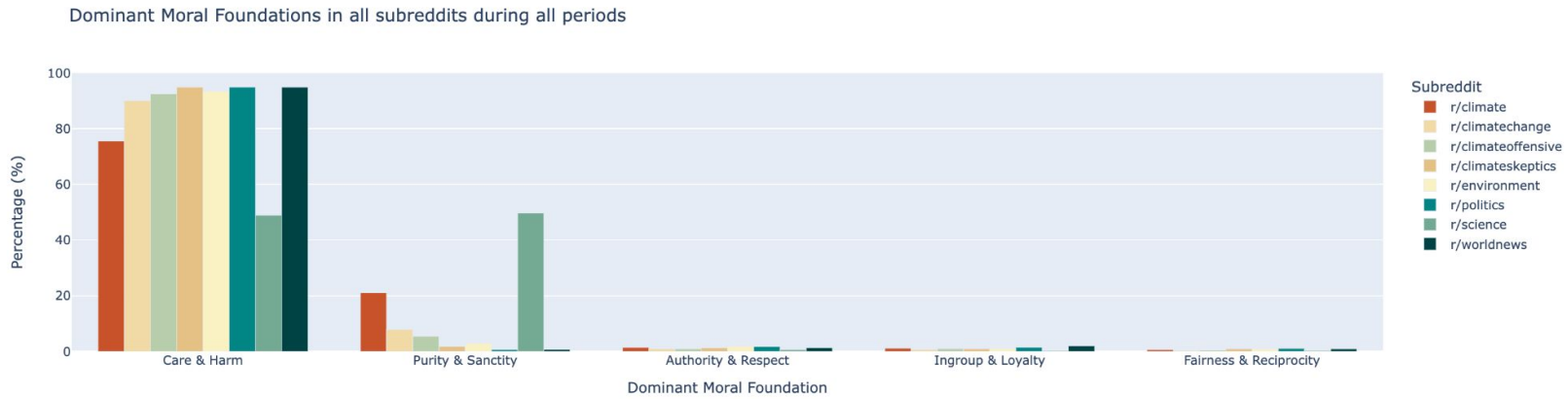
Moral foundation theory argues that there are five basic moral foundations: (1) harm/care, (2) fairness/reciprocity, (3) ingroup/loyalty, (4) authority/respect, and (5) purity/sanctity. These five foundations comprise the building blocks of morality, regardless of the culture. The following analyzes data from 8 subreddits related to climate based on the dominant moral foundations. Major weather events tend to increase discussion of climate online. To inspect the effect of this increased discourse, we analyze reddit comments before, during, and after Hurricane Ian (September 23, 2022 – September 30, 2022).

Select a subreddit: All Subreddits

Select a time period: All Periods

Frequency of Dominant Moral Foundation

Comments were classified by their dominant moral foundation by taking the cosine similarity between each comment and words in an Expanded Moral Foundation Dictionary. Each comment is classified to the moral foundation with which it is most aligned. The percentages reflect the percent of comments in that subreddit classified to that moral foundation.




Questions?