From Ranking to Clustering Cultures

Towards a time-varying network analysis of YouTube algorithms on Global Warming and Climate Change

FACILITATOR João Guilherme Bastos dos Santos

Carlos Eduardo Barros

Tal Cohen

Alessandra Facchin

Julia Jasińska

Luis Landa

Helena Movchan

Michal Salamon

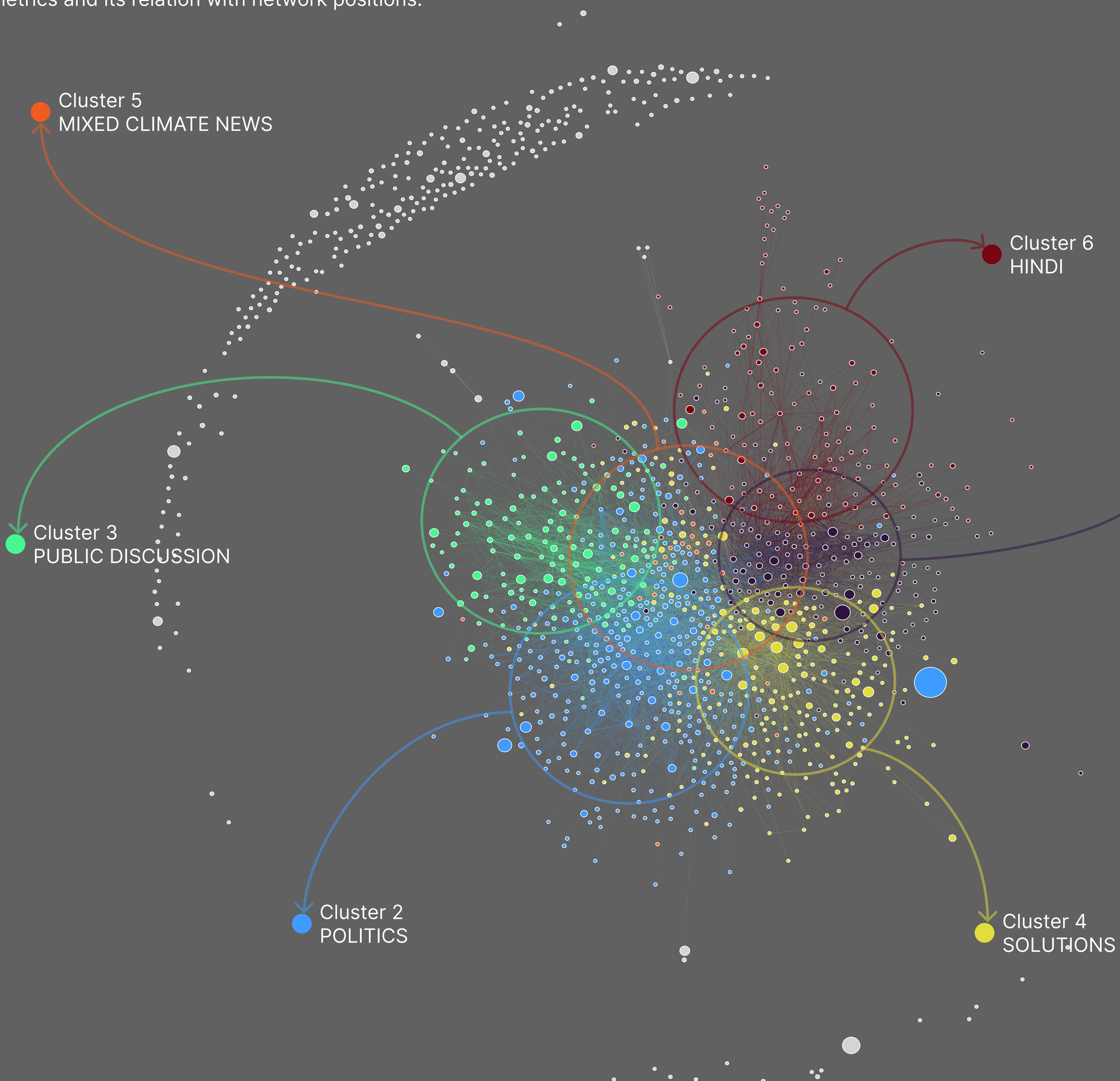
Aleksy Szymkiewicz

Xiaoyue Yan

PARTICIPANTS

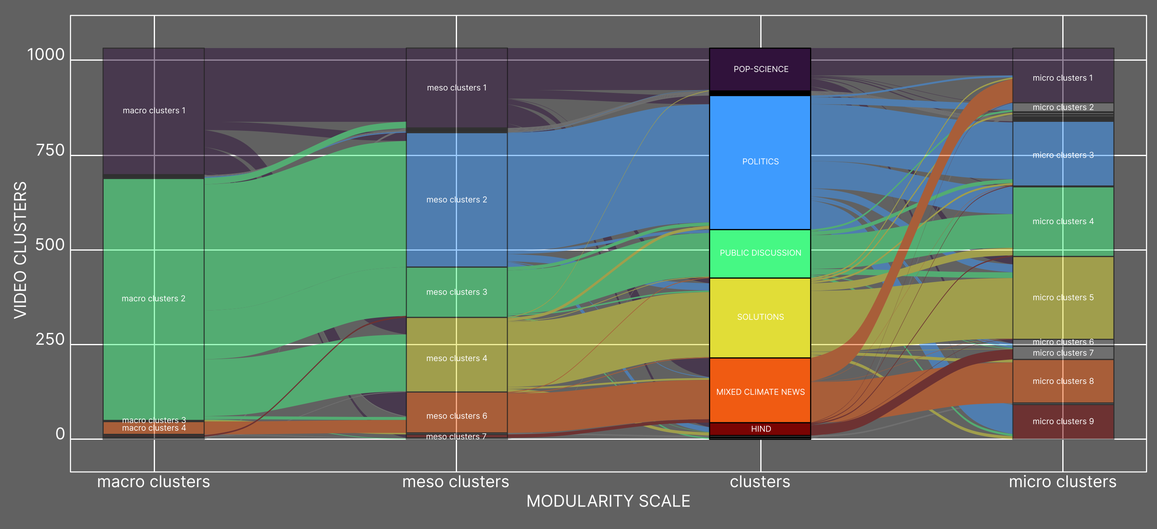
NETWORK ANALYSIS

Network analysis encompasses Louvain clustering for grouping videos networked by related videos recommendations, time-varying PageRank centrality to identify the most central videos and clusters on a daily basis, recommendation flows inside the network, engagement metrics and its relation with network positions.

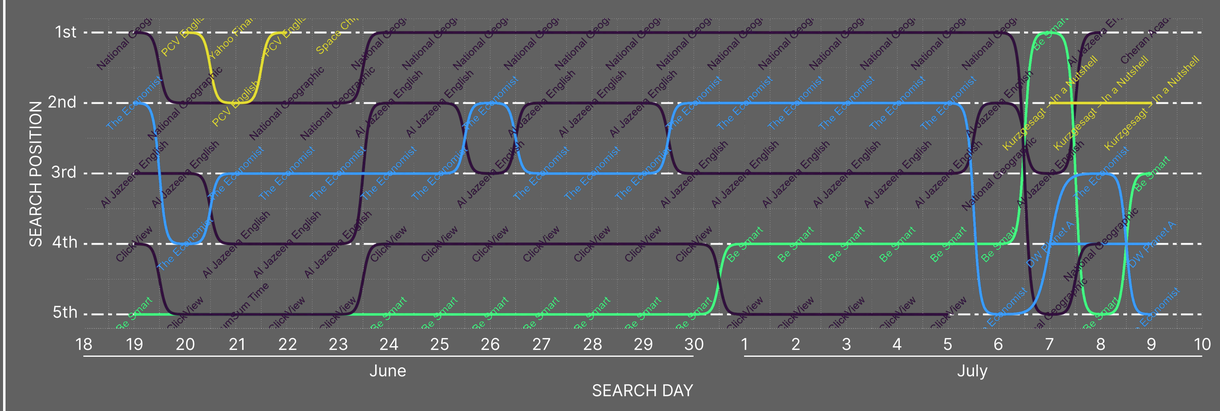


CLUSTERS FRAGMENTATION ACCORDING TO MODULARITY ADJUSTMENT

The search rankflow shows the disturbance in the steady patterns found from June 23rd to the 5th of July, possibly because July 4th was the hottest day on record. The cluster about Solutions went back to the top five in the second position for three consecutive days and the Public Discussion surpassed Pop-science and Politics clusters to go from fourth to the first position.



SEARCH RANK FLOW



The search rankflow shows the disturbance in the steady patterns found from June 23rd to the 5th of July, possibly because July 4th was the hottest day on record. The cluster about Solutions went back to the top five in the second position for three consecutive days and the Public Discussion surpassed Pop-science and Politics clusters to go from fourth to the first position.

From Ranking to Clustering Cultures

Towards a time-varying network analysis of YouTube algorithms on Global Warming and Climate Change

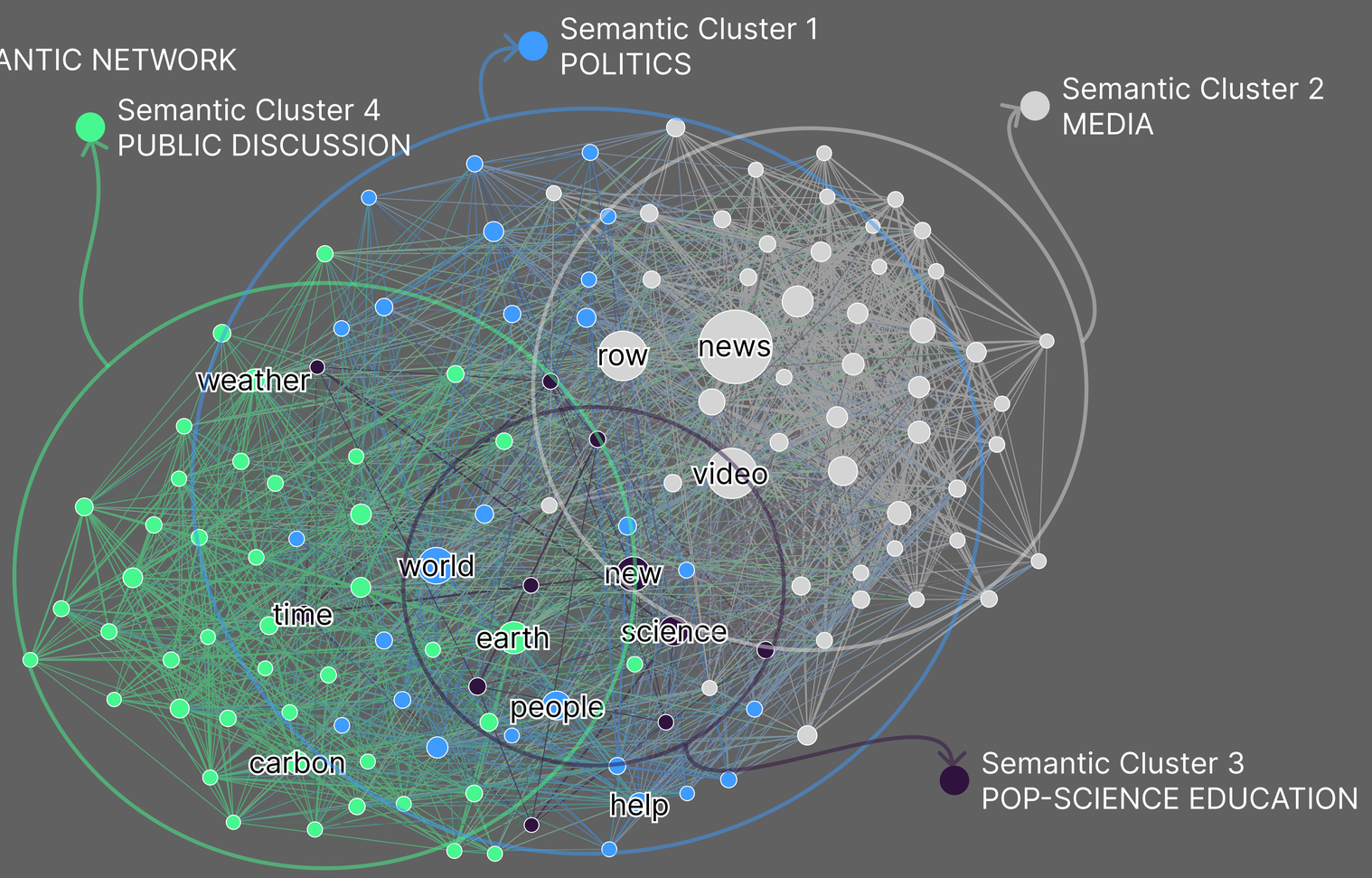
TEXT ANALYSIS

Text analysis encompasses tf-idf of video descriptions and subtitles - when available - to check if clusters correspond to specific themes or perspectives, topic modelling, semantic networks and Reinert algorithm to check the main issues discussed in the corpus and correspondence analysis to understand the position of each video inside the debate.

WORDTREE

This word tree aggregates and visually represents all of the obtainable transcripts of cluster 3 videos. It allows us to explore how specific phrases, such as “climate change is,” are used across the corpus. Upon querying this phrase, we can trace the branches of the tree to see the most common continuations of the statement. In our case, the content that follows “climate change is” appears to be predominantly denialist or conspiritual in nature, indicating such viewpoints are prevalent within the analyzed corpus. This visualization tool thus provides valuable insights into the dominant discourse patterns related to climate change in our dataset.

SEMANTIC NETWORK

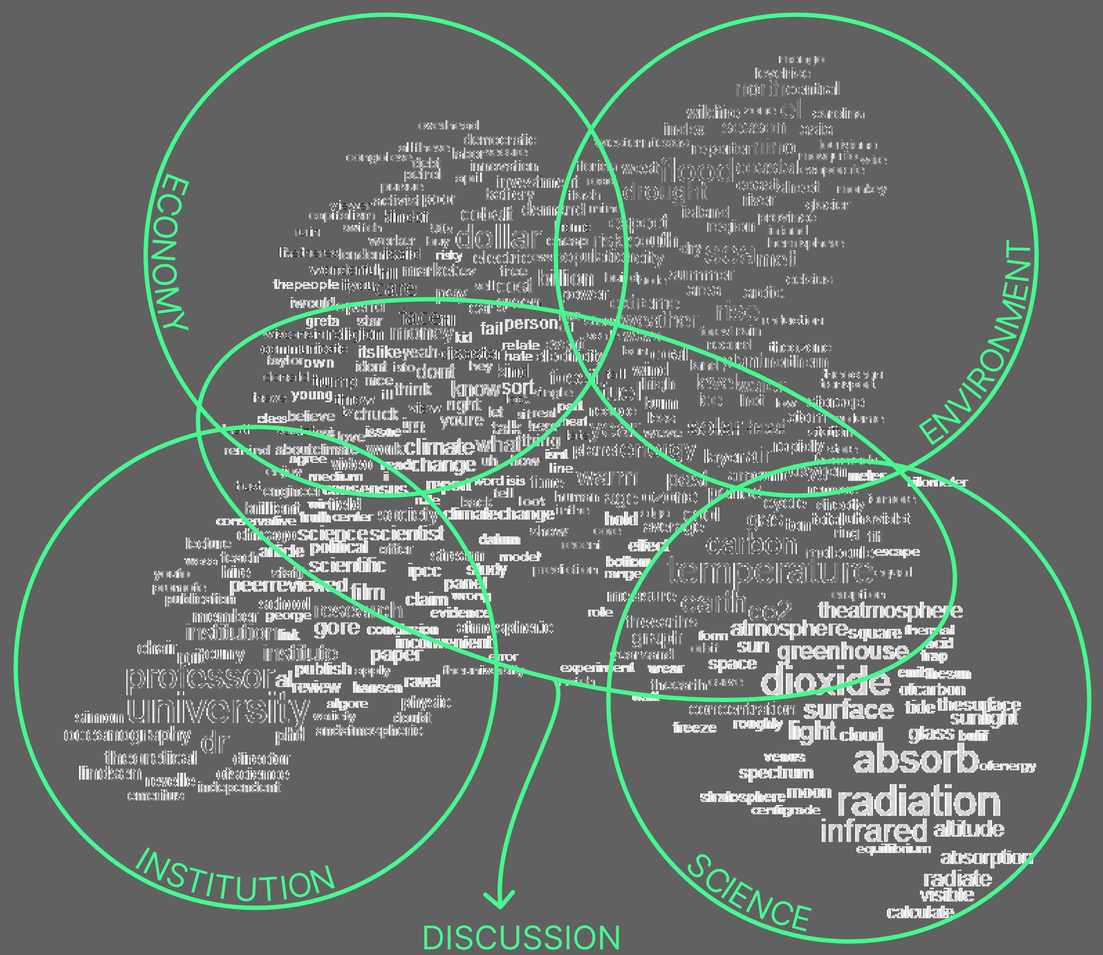


This study examines the discourse network surrounding climate change and global warming by analyzing YouTube video descriptions. The network comprises 130 actors categorized into four clusters, representing distinct frameworks that can be interpreted as follows: the problem itself, pop-scientific evaluation, communication medium, and potential solutions. The network highlights actors with the highest betweenness centrality scores, serving as labels within the network visualization.

Cluster 1 emphasizes the significance of understanding the problem and facilitating knowledge exchange among nations. **Cluster 2** represents the medium through which the discourse on global warming is disseminated. It encompasses various social networks, traditional news outlets, and media channels. **Cluster 3** characterizes the popular-science perspective on the phenomenon, providing accessible information and insights. It frames the problem as a scientific issue, necessitating an understanding of context, history, and empirical data. **Cluster 4** showcases the direct terminology used to describe global warming and the climate crisis.

WORDS IN CLUSTER

The most interesting cluster - it represents the public debate on climate change.It contains both content from reliable sources and misleading material.Analysing the transcripts allowed us to identify the themes raised: scientific discourse, green politics (issues related to economics, regulation, planning and investment) and issues related to the impacts of climate change (both environmental and political, which are part of the green political narrative).Issues raised by the fake news sub-cluster include undermining the institution of science, economic issues and rising temperatures.



PARTICIPANTS Xiaoyue Yan Aleksy Szymkiewicz Michal Salamon Helena Movchan Luis Landa Julia Jasińska Alessandra Facchin Tal Cohen Carlos Eduardo Barros FACILITATOR João Guilherme Bastos dos Santos

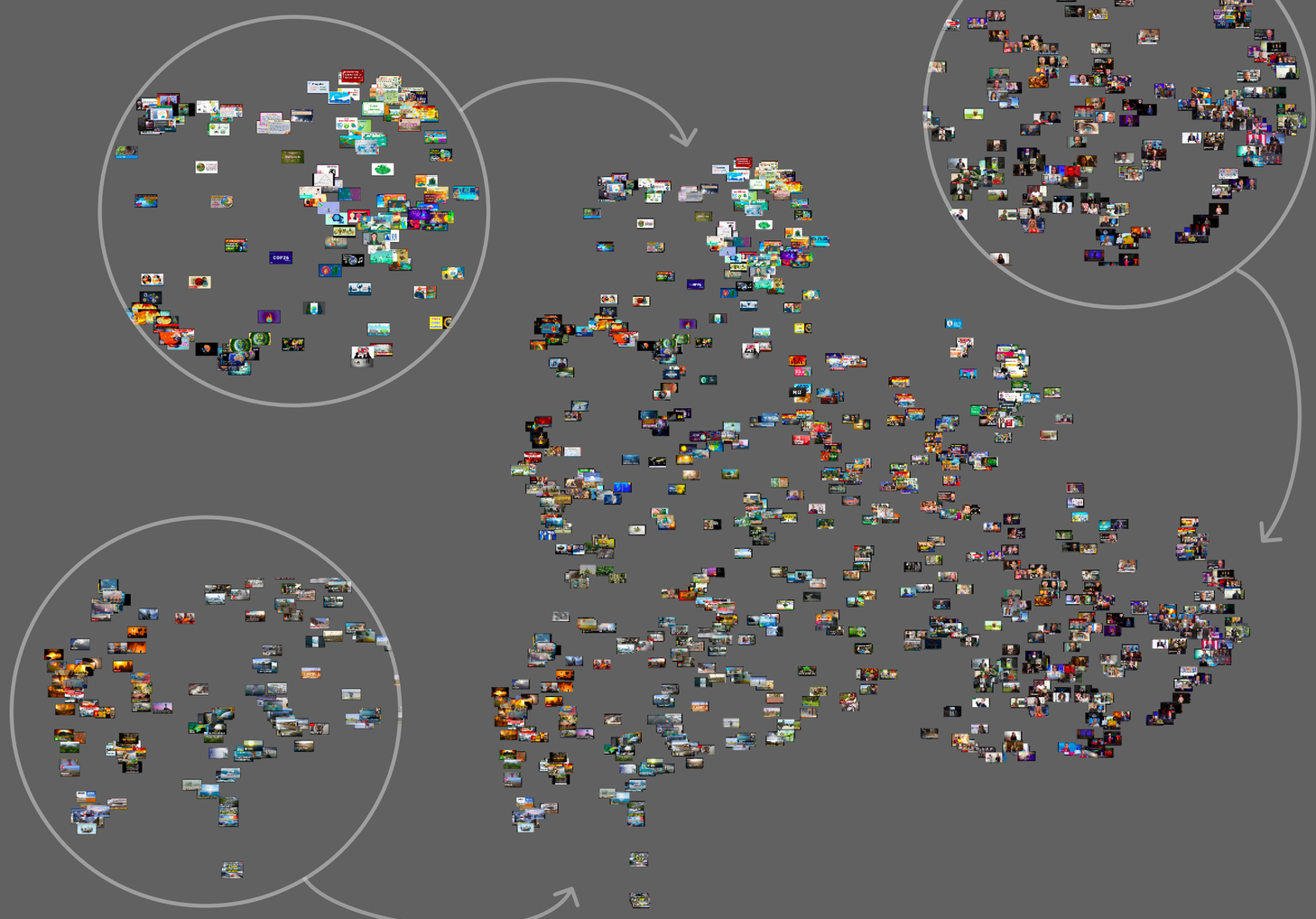
From Ranking to Clustering Cultures

Towards a time-varying network analysis of YouTube algorithms on Global Warming and Climate Change

João Guilherme Bastos dos Santos FACILITATOR
Carlos Eduardo Barros
Tal Cohen
Alessandra Facchin
Julia Jasińska
Luis Landa
Helena Movchan
Michal Salamon
Aleksy Szymkiewicz
Xiaoyue Yan
PARTICIPANTS

IMAGE ANALYSIS

Image analysis encompasses video thumbnails' images to be analysed using PixPlot image sorting, cluster's specific folders to be analyzed qualitatively to check if clusters correspond to specific grammars, aesthetics or symbols.



The illustration presents the Umap which resulted from analyzing six semantic clusters using PixPlot. The concentration of images happens in three vertices, with even distribution of various images in the center. The interpretation of the distribution of semantic clusters and PixPlot clusters can be conducted by observing the placement of images on the map.

PixPlot created 10 clusters, all of which have their own characteristics and properties. The similarities between them result in three points of concentration, separating all images into graphs, photographs of people, and photographs of nature. Furthermore, images can be grouped by their original semantic clusters, which allows the analysis of dominant visualities and formats in each cluster (Fig. X). It allows for the calculation of images assigned to each PixPlot cluster within the semantic cluster, and establishing the dominating visual style.

MULTIMODAL ANALYSIS

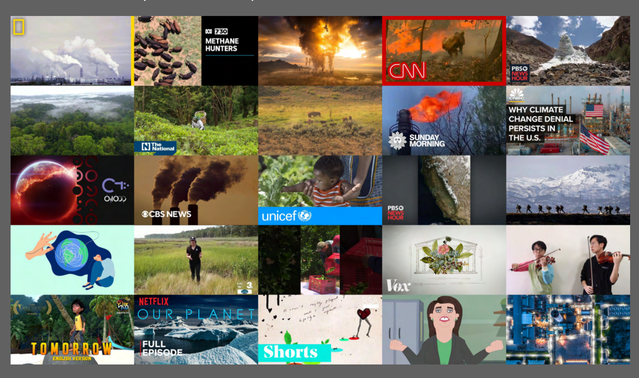
Multimodal clustering is the process of grouping data that consists of different types of information (e.g., text, images, audio, etc.) into meaningful clusters based on their similarities and relationships.With the model clip-ViT-B-32 from sentence transformer, a multimodal clustering was carried out using BERTopic on the video description and video thumbnail data.

TOPIC 0

Video description: the cause and solutions of climate change
Thumbnail description: often contain texts in a large font size with astonishing expressions

TOPIC 1

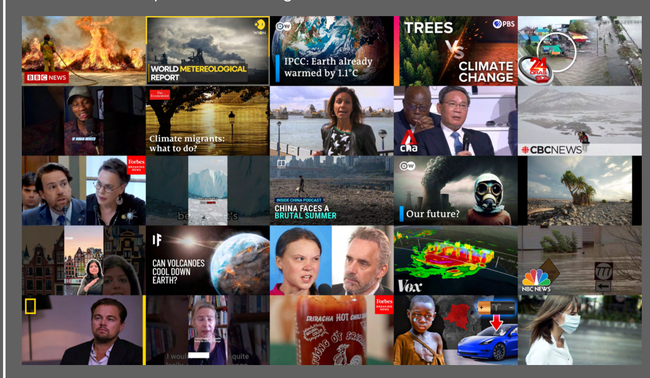
Video description: introducing climate change and "hard" facts about the issue
Thumbnail description: landscapes and animals



When it comes to the temporal change of the topics within the data collection period. One particular example is the third topic, the influence of climate change. An increase in the number of collected videos was observed since 7th of July, likely due to the hottest day in the world on 4th of July, and "heatwaves" became the keyword of the cluster for the first time on 9th of July.

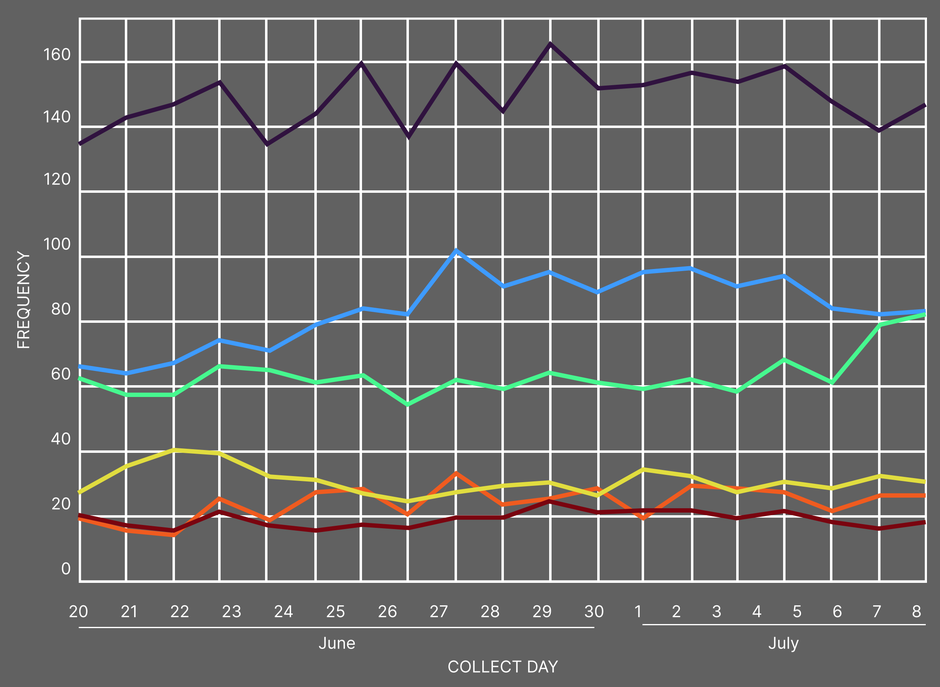
TOPIC 3

Video description: India-related climate change news in Hindi
Thumbnail description: contain English texts



TOPIC 2

Video description: the influence of climate change
Thumbnail description: a person reporting the issues



TOPIC 0 TOPIC 1 TOPIC 2 TOPIC 3 TOPIC 4 TOPIC 5

TOPIC 4

Video description: India-related climate change news in English
Thumbnail description: contain Hindi texts

TOPIC 5

Video description: international climate change news in a diverse range of languages
Thumbnail description: a mixture of landscape, human reporter, and texts