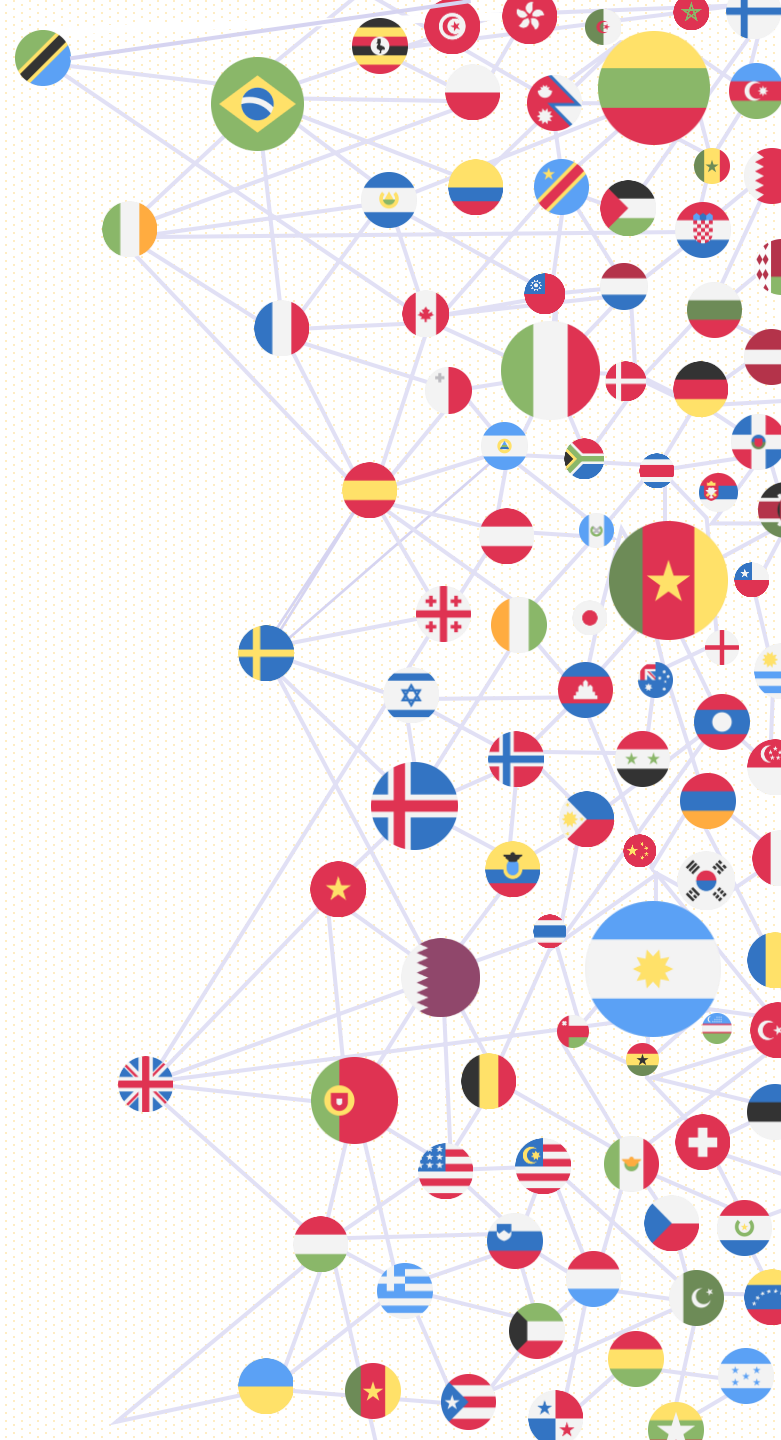


National anthems analysis

- *Jakub Koral*
- *Marcin Kostrzewa*
- *Marcin Miśkiewicz*



Agenda

1. Interesting facts
2. NLP analysis
3. Cluster analysis

The source of data

We analyzed the English translations of the anthems, scraped from the website

<https://nationalanthems.info/>

licensed under the

[Creative Commons Attribution 4.0 International](#) 



Which country's anthem is
the longest?

Longest anthem



Netherlands
601 words

Shortest anthem



Jordan
18 words



**Bosnia and
Herzegovina**



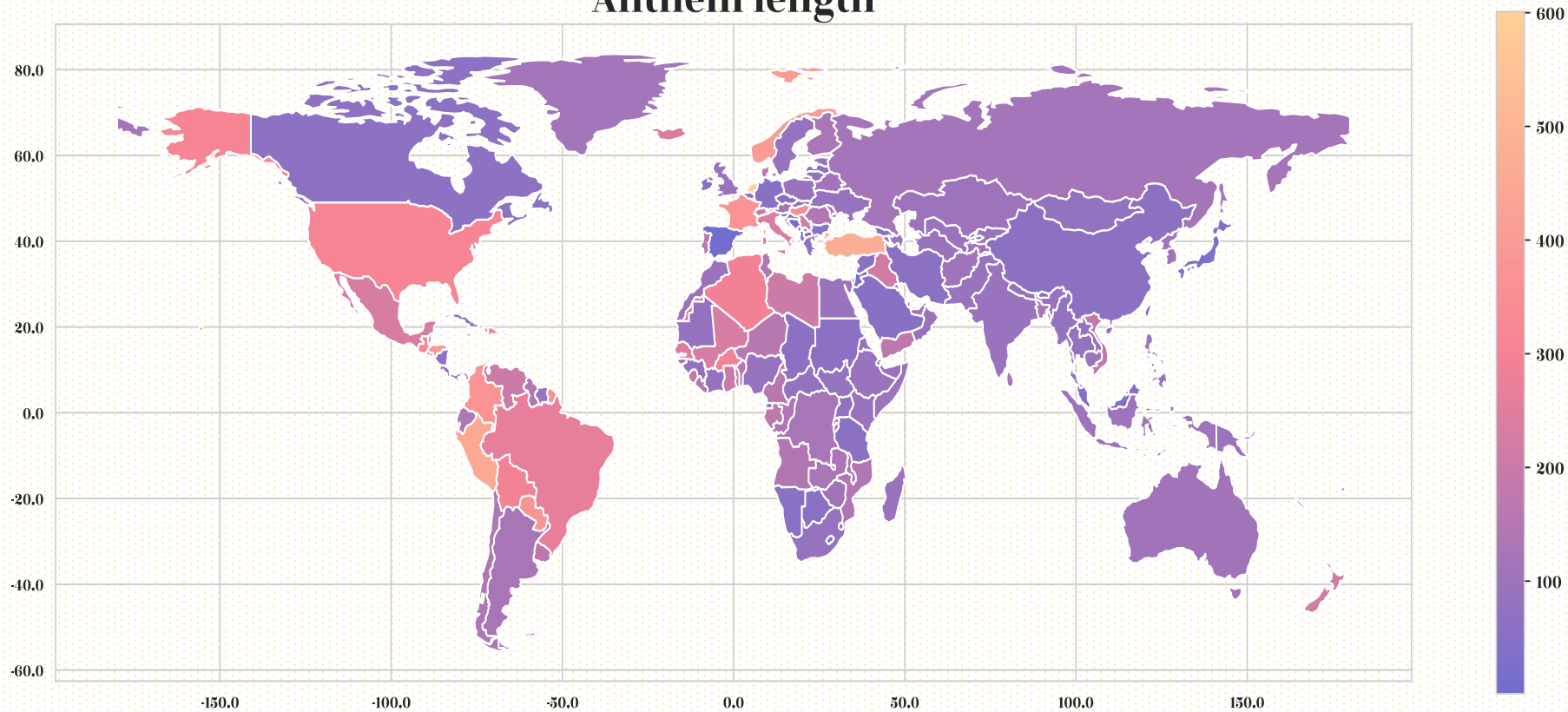
Kosovo



Spain

No official lyrics

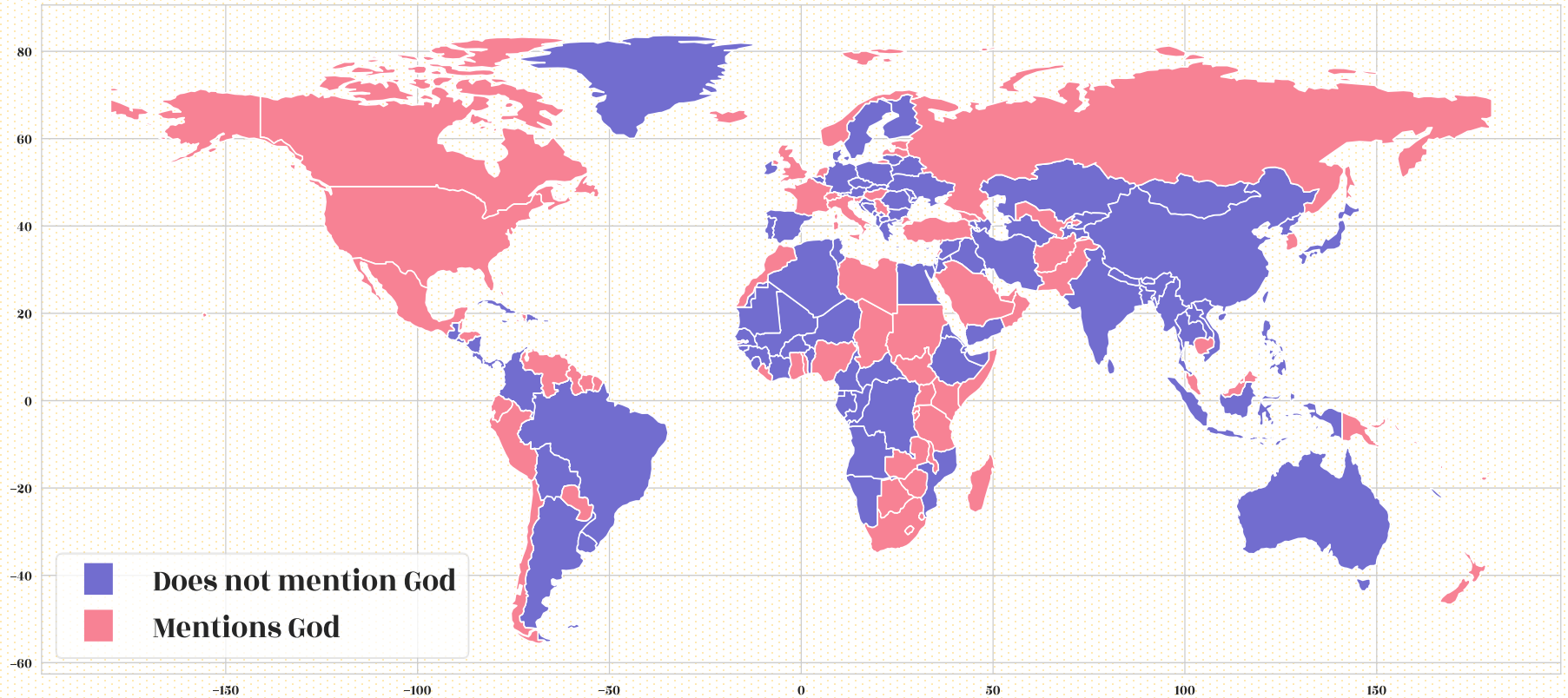
Anthem length



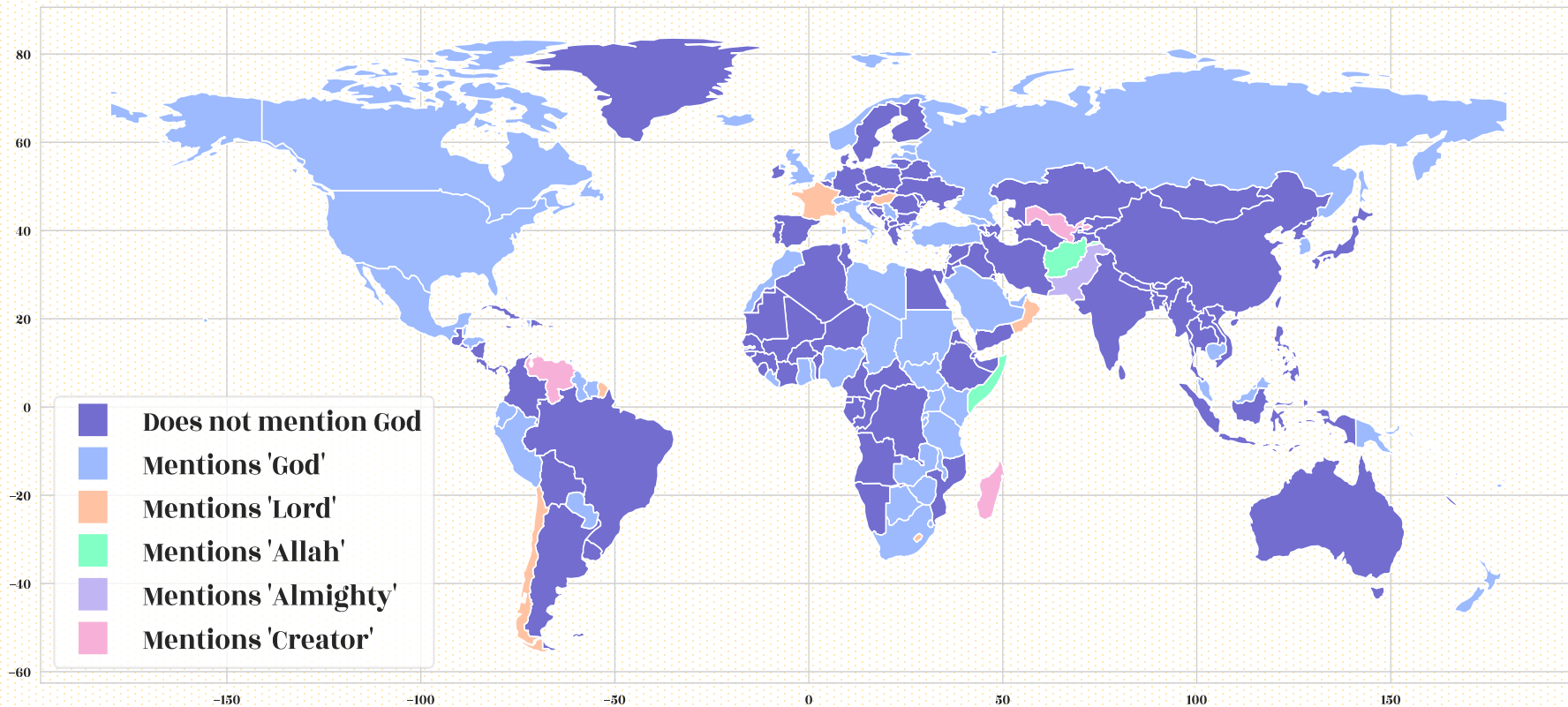


Which national anthems mention God?

Mentions of God in national anthems



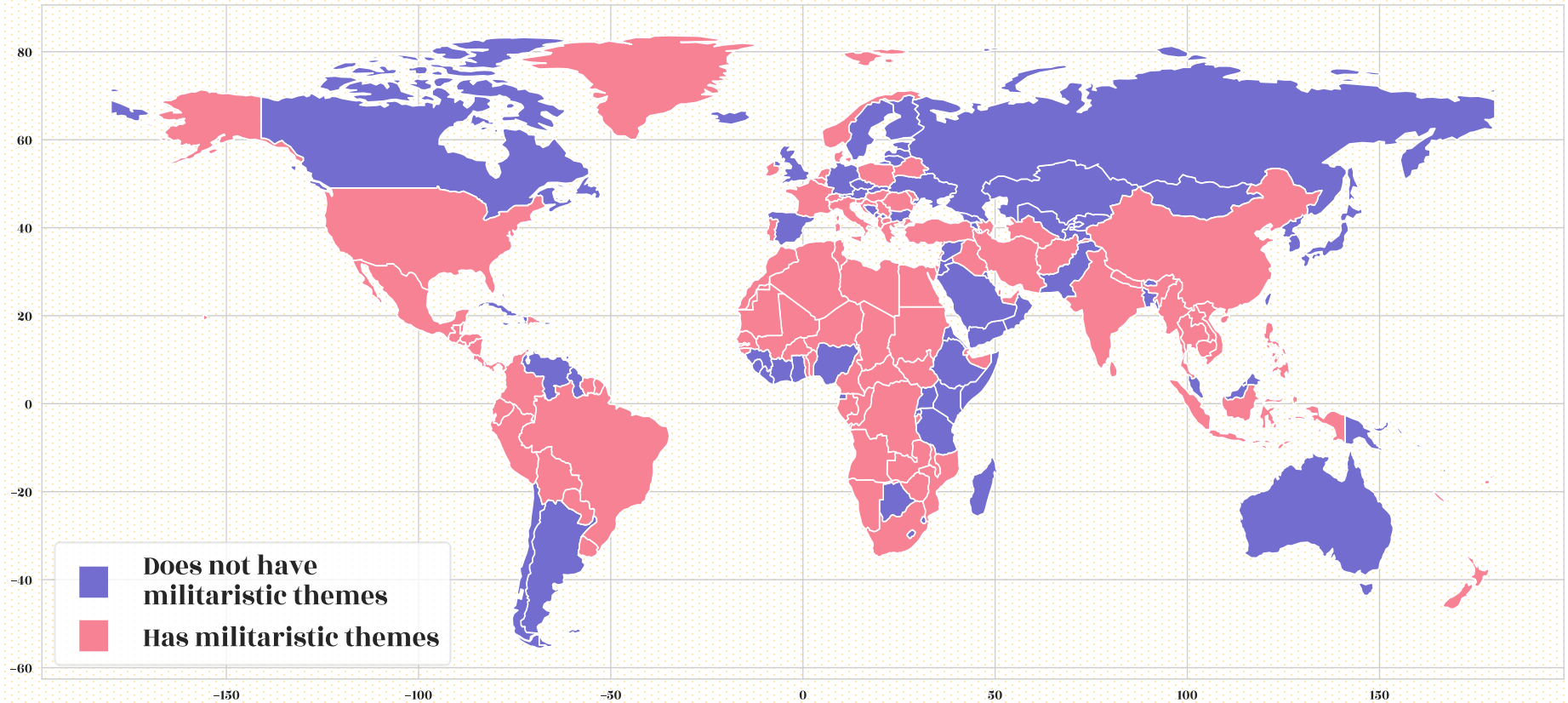
Mentions of God in national anthems: further subdivisions





Which national anthems have
militaristic
themes?

Militaristic themes in national anthems





What animals are mentioned in national anthems?

Senegal



Lebanon



Dominican Rep.



Libya



Zambia



Italy



Guatemala



Qatar



Ecuador

Honduras



Paraguay



Falkland Is.






France














What are the most common

n-grams & collocations?

Word	Frequency
 land	247
 let	173
 god	168
country	134
free	122
people	117
thy	109
love	104
oh	98
glory	97

Bigram	Frequency
 long live	27
 god bless	23
 native land	17
free land	11
thousand years	11
oh god	11
witness witness	10
land free	10

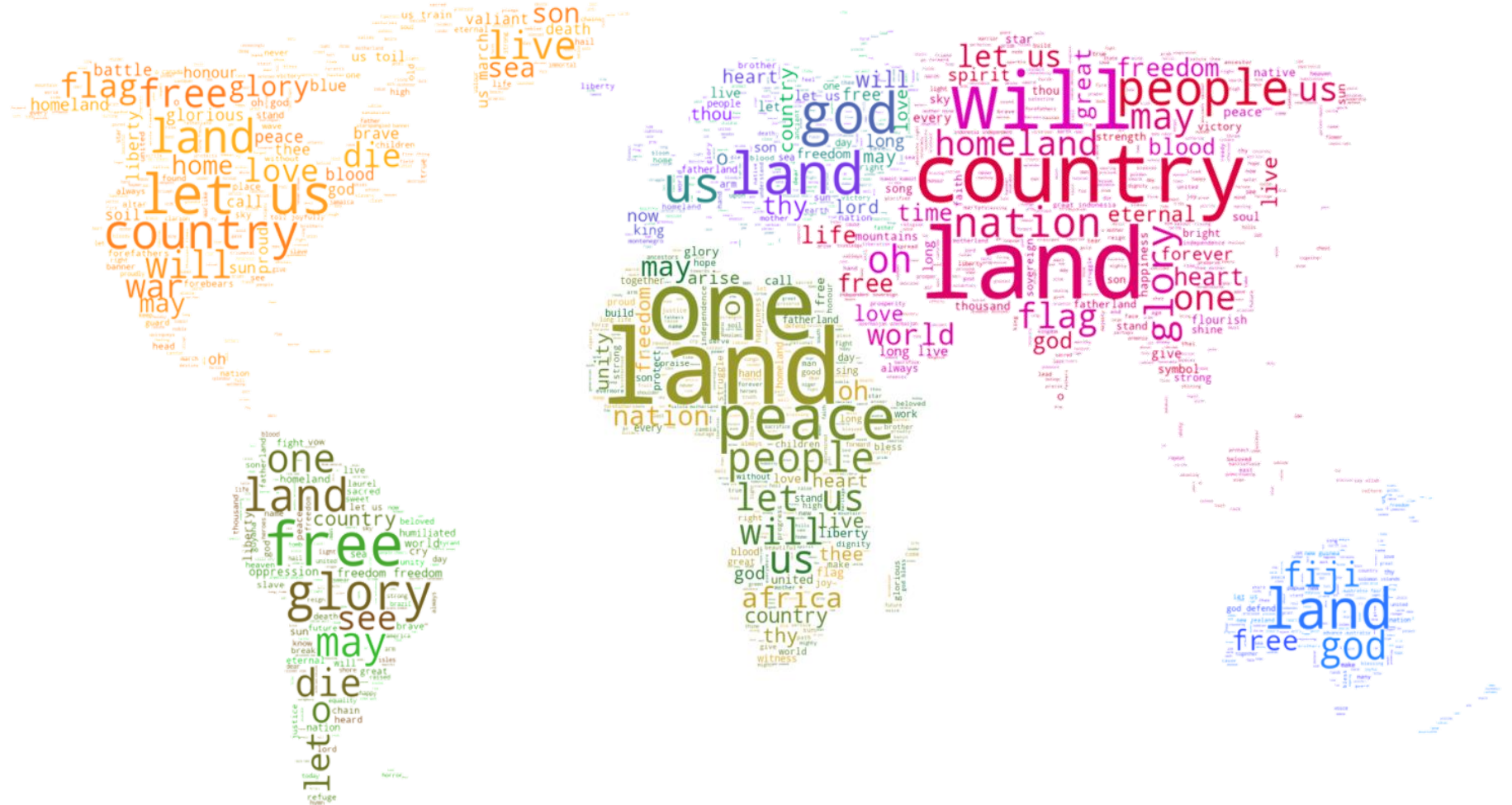
Bigram collocation	Frequency
 let us	127
 long live	27
 god bless	24
may god	20
native land	17
thousand years	13
give us	12
oh god	11

Trigram collocation	Frequency
 long life peace	27
 let us march	23
 let us unite	17
let us train	13
let us toil	11
ancient valour rise	11
let us swear	11



What are the most common
words by
continent?

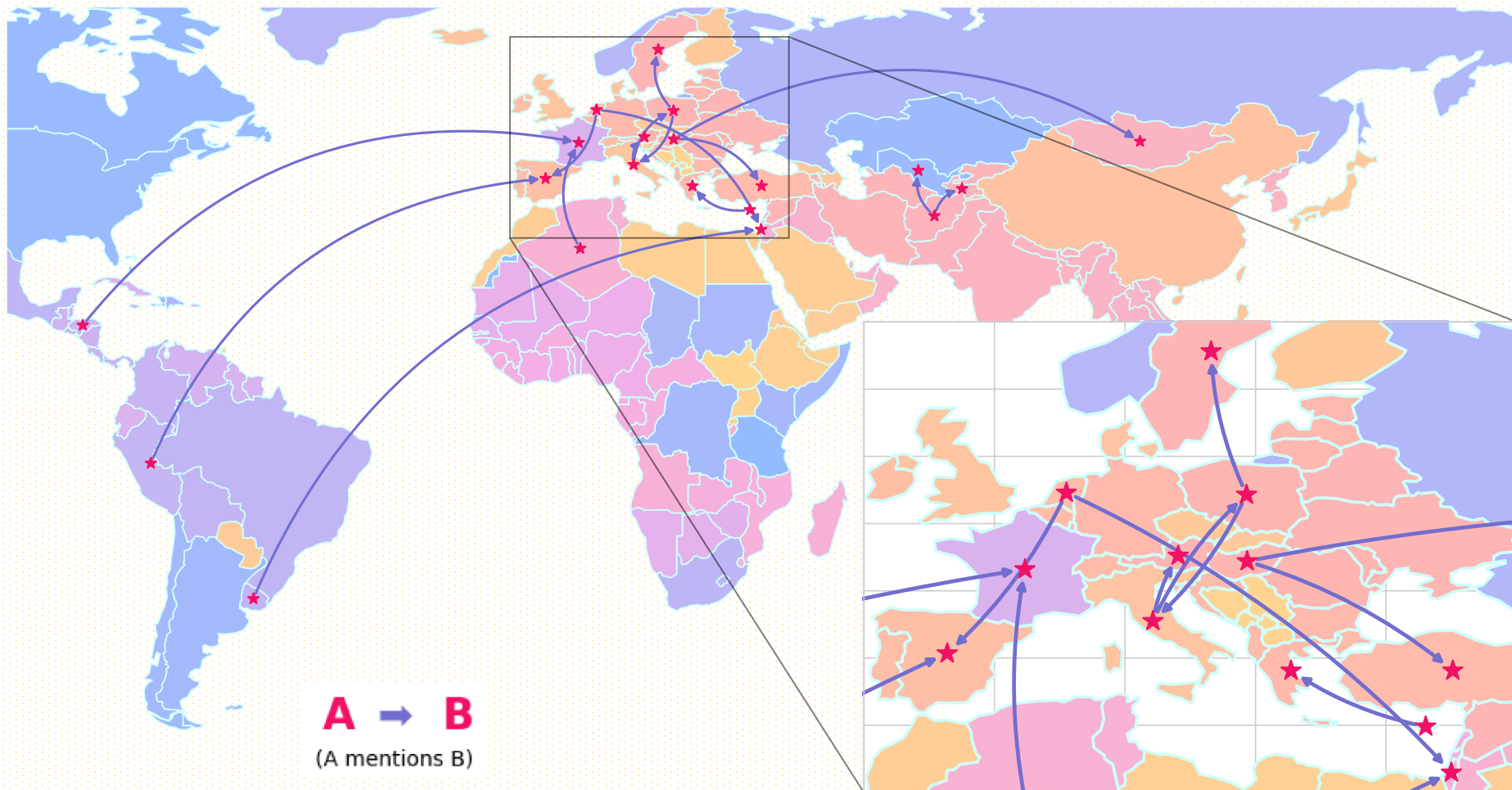
The most common words in national anthems by continent





Which countries
mention
each other?

Direct mentions of other countries in national anthems



Mentions

- Peru → Spain,
- Honduras → France,
- Algeria → France,
- Poland → Sweden, Italy,
- Hungary → Turkey, Mongolia,
- Netherlands → Israel, Spain,
- Italy → Poland, Austria,
- Cyprus → Greece,
- Uruguay → Israel
- Afghanistan → Tajikistan, Uzbekistan,

Some cases had to be removed manually...

National Anthem of Lebanon

“(...) Throughout the world her good deeds flow from pole to pole.”



National Anthem of Honduras

“(...) Like an Indian maiden you have been sleeping, Lulled by the resonant song of your seas,.”





Clustering

Clustering methodology

Frequency based



TF-IDF

1010 1010
1010 1010
1010 1010
1010 1010

K-Means

Context based



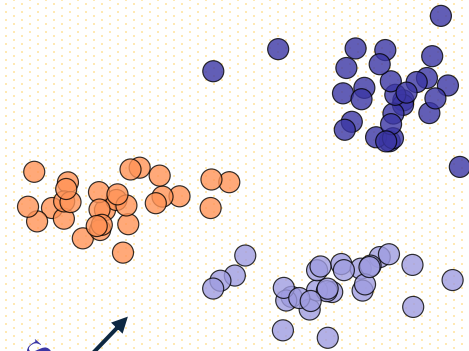
1010 1010
1010 1010
1010 1010
1010 1010

umap

1010
1010
1010
1010

AGNES

Clusters





Frequency based clustering

Cluster 1



Cluster 2



Cluster 3



Cluster 4



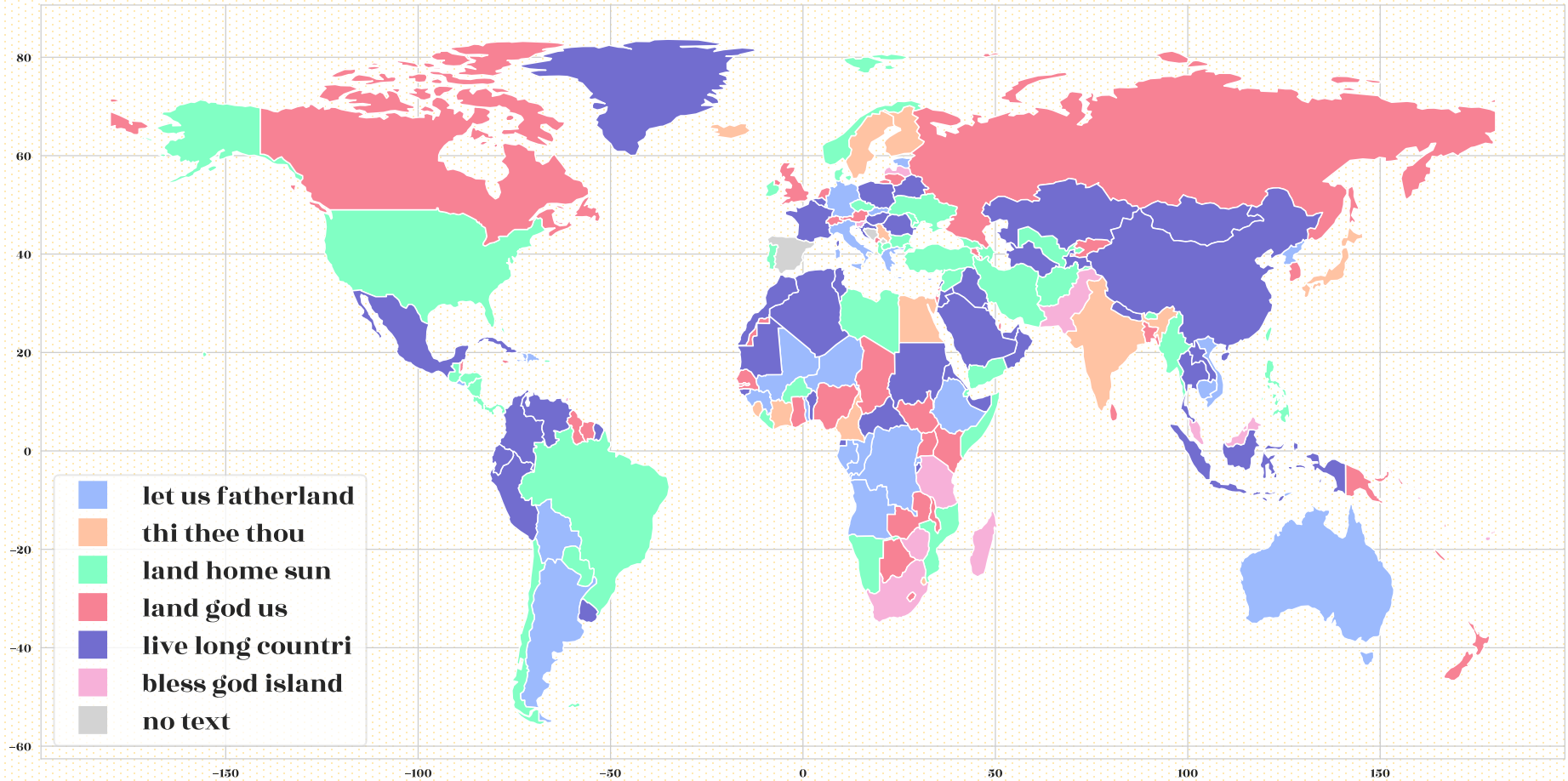
Cluster 5



Cluster 6



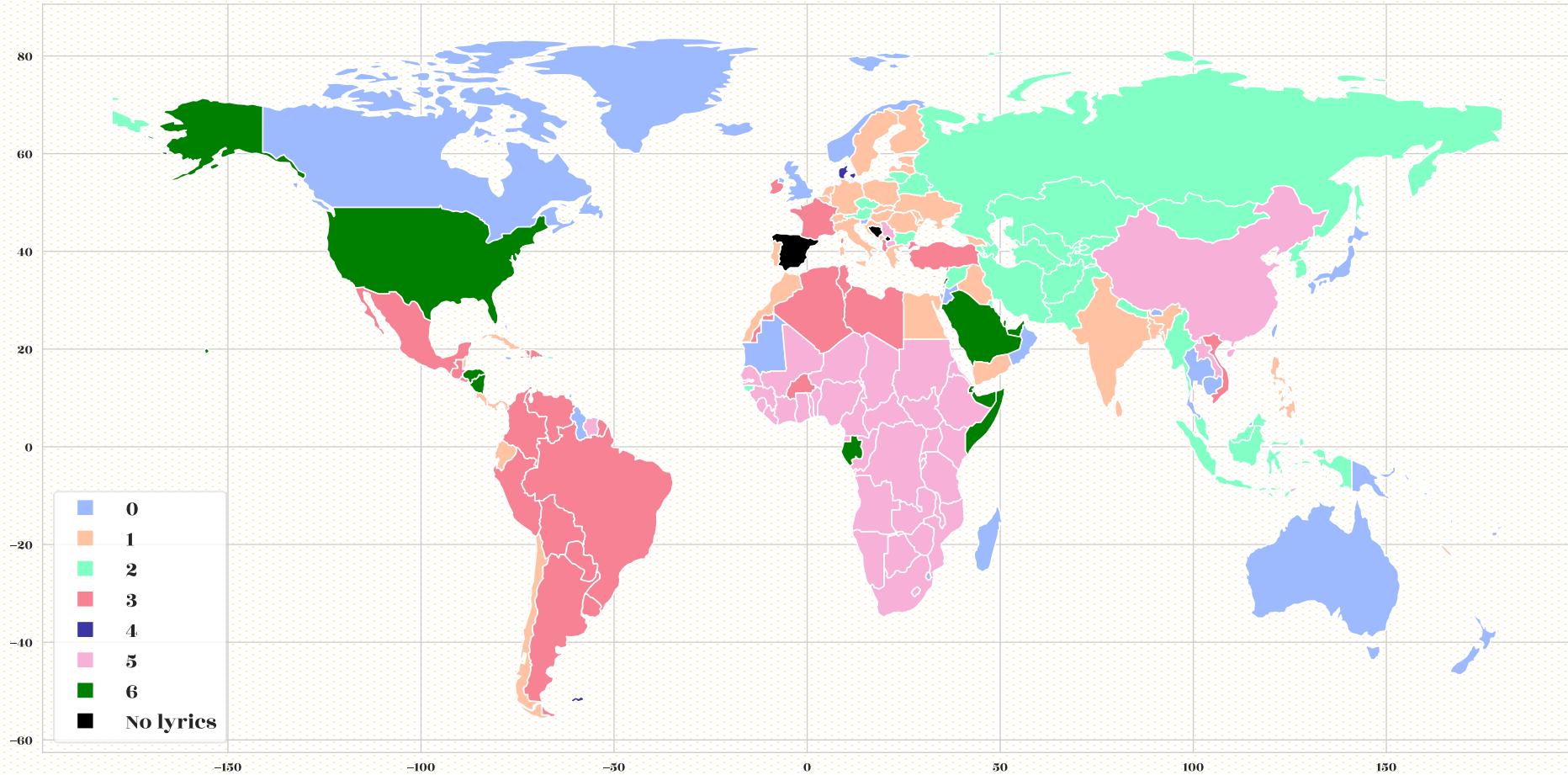
Clustered anthems: K-Means





Context based clustering

Clustered anthems: AGNES



How to explain these results?

**Let's ask an expert —
GPT-4.**

Exemplary GPT-4 trials

- **Cluster 0 theme:** *All of these are the English translations of the national anthems of different countries.*
- **Cluster 1 theme:** *All of these national anthems belong to countries that are part of the Commonwealth of Nations.*
- **Cluster 3 theme:** *All of these national anthems emphasize the themes of freedom, national pride, resistance against oppression, and unity*
- **Cluster 5 theme:** *All of these national anthems express a strong sense of national pride, unity, and hope for the future. They also all mention the desire for peace, freedom, and prosperity for their respective nations.*

Conclusions

1. Majority of anthems include military themes. Animals are relatively rarely mentioned.
2. There is no one definite criterion (geographical, historical, religious), that is responsible for the clustering of anthems.

Possible explanation (?)

Anthems were created in a similar time frame (19th and 20th centuries), when the impact of globalisation could be significant.

3. Themes present in anthems seem to be universal across the whole world. They include: invocation to deity, calls for action and unity and appreciation of the country's past.