

# THE WORLD ACCORDING TO TIKTOK

## MAPPING DIGITAL BOUNDARIES AND ALGORITHMIC PROXIMITIES

### PROJECT DESCRIPTION 1/2

TikTok is one of the most popular, and potentially influential, digital platforms. The platform's main feature is its For You Page (FYP), a continuous feed of content that is algorithmically curated by the platform's recommendation system (TikTok 2020). This recommendation system is based on a combination of factors, such as user interests and interactions, video information, device and account settings, as well as the device and user's geographical location. This project studies the FYP from a cross-national perspective, by investigating emerging patterns, trends, concerns, and thematic clusters when (non-authenticated) users access TikTok in different countries. What algorithmically-mediated cross-national boundaries and proximities can we map out?

### RESEARCH QUESTIONS

We aim to map networks, discourses, trends, and key actors promoted by TikTok in a cross-national analysis to understand how TikTok remediates proximities and boundaries among geographic regions. Further, we intend to show how TikTok potentially navigates controversial and/or contested issues in said regions. We ask:

1. How can we map algorithmic proximity and digital boundaries between countries?
2. How are discourses, languages and audio algorithmically represented on a cross-national level?
3. How are human rights issues algorithmically represented in different geographical and cultural contexts?

### DATA COLLECTION METHODOLOGY

In preparation, Tracking Exposed collected two separate datasets containing data from summer and winter of 2022. Each dataset consists of the first 8 videos that appear on the FYP of a non-logged-in TikTok user across 197 countries, scraped up to 4 times a day. Both datasets consist of roughly 300.000 recommended videos, in each of which 47.000 are unique. This data collection relies on the Residential IPs Network developed by BrightData. Methodologically, we used a combination of exploratory data analysis, manual and semi-automated content analysis, network analysis/visualization to get a complex representational overview of "the world according to TikTok", based on patterns of algorithmic proximity.

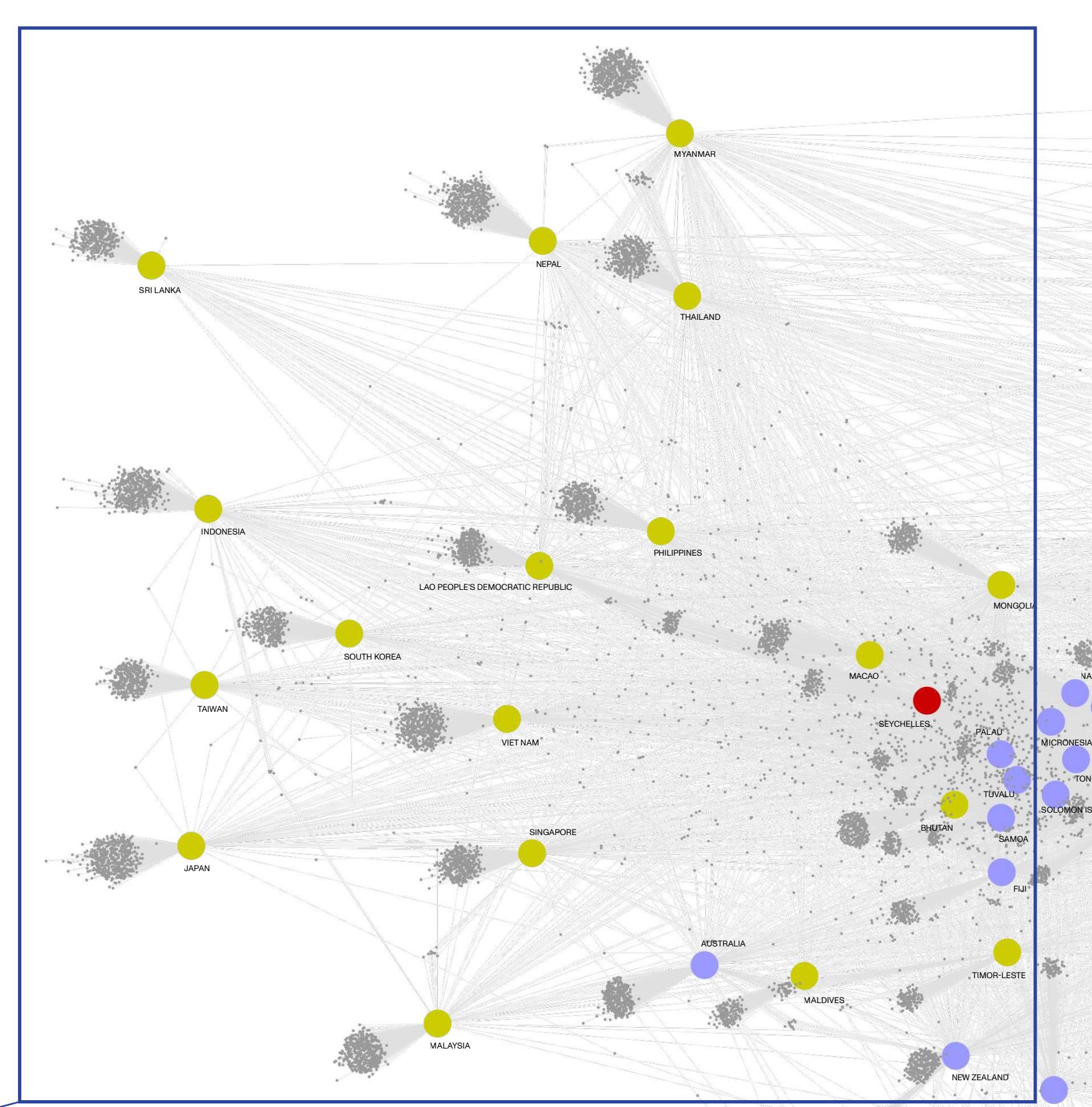
### DIGITAL BORDERS ON TIKTOK

In examining the macro distribution of videos on TikTok, we were able to sketch a 'world-according-to-TikTok' map. The motivation for this task was to determine whether the diffusion of videos across different countries follows a parallel structure to political and geographic boundaries. To do so, we created a map using 'videoID' metadata collected from each country to visualize the division of video content. The graph shown here is from the data gathered during the Winter season—differences between the summer and winter graphs were negligible at best. As such we chose to present the more recent data due to its timeliness.

Our findings show that the videos indeed follow physical, political boundaries. The clusters formed demonstrate a parallel with geographical proximity.

#### LEGEND

GEOGRAPHIC AREA
North & South America
Europe
Sub-Saharan Africa
Middle East and North Africa
Asia
Pacific

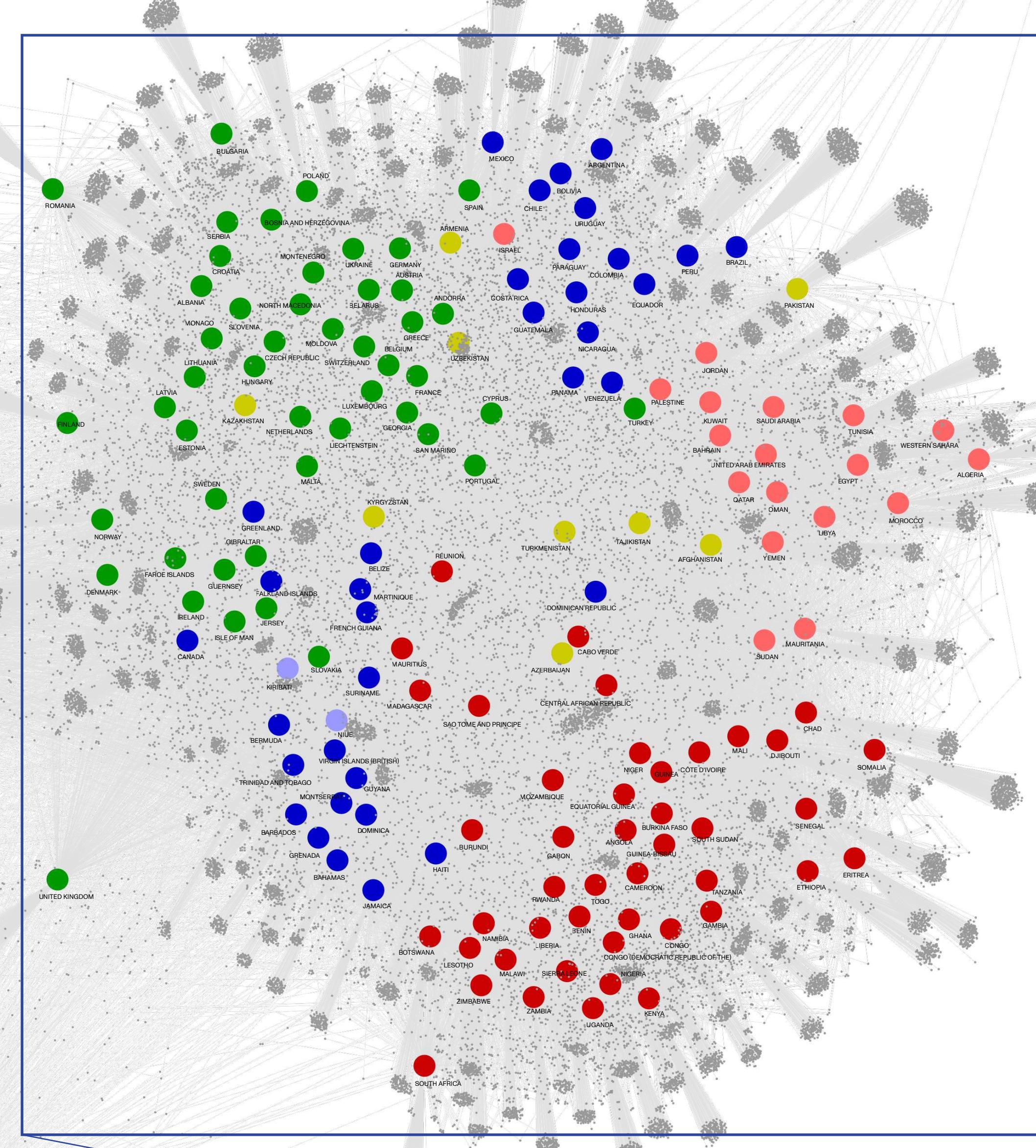


#### FINDING

East and South East Asian countries seem to have very insular TikTok feeds.

#### FINDING

Bangladesh is among a rare few outliers in the dataset. In Q3 of 2022, Bangladesh removed just under 6 million videos from the platform.

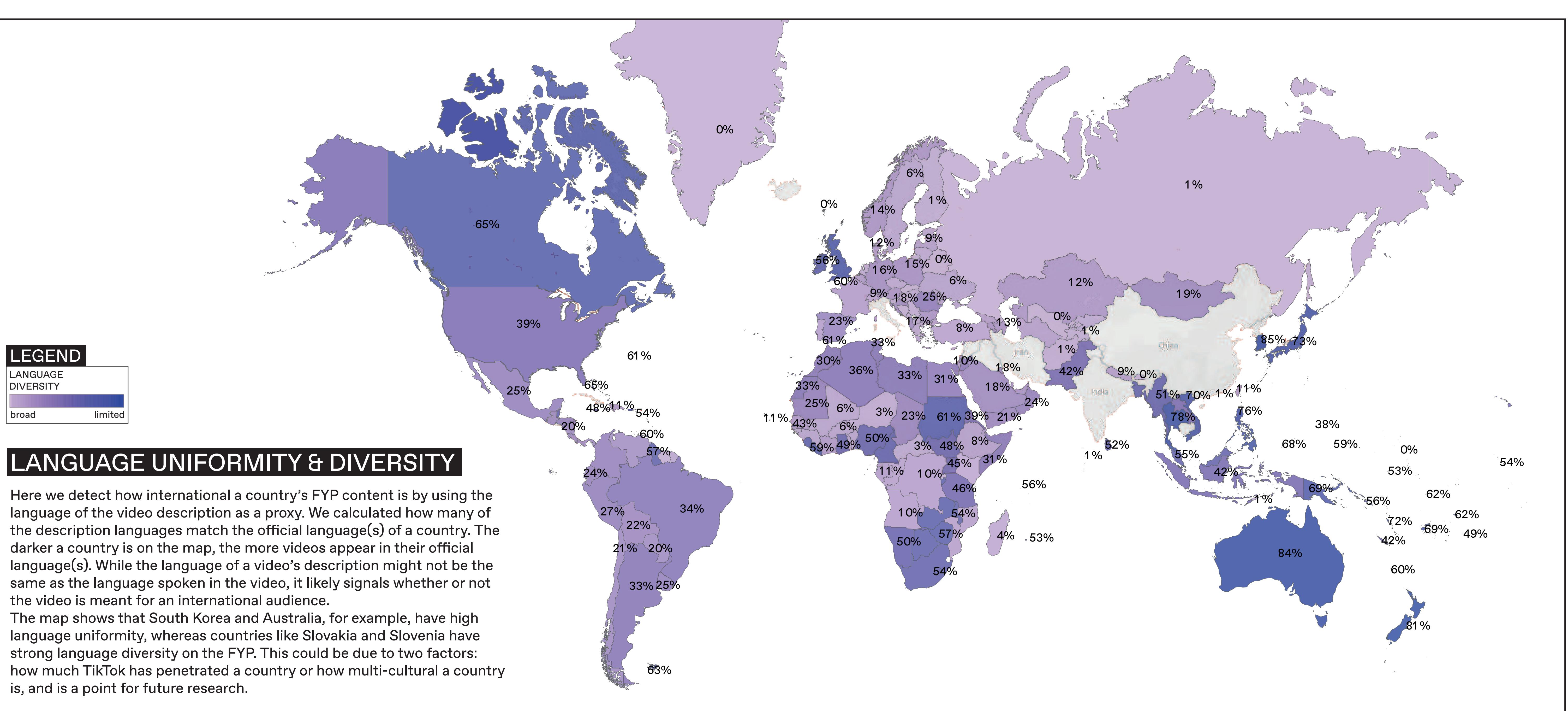


#### FINDING

The United States and its territories have a common feed separate from the rest of the world. Likewise, Russia has a feed unique to itself.

#### FINDING

Most countries form a major conglomeration, that nonetheless cluster along geographic lines. Other factors also seem to play a part, such as Spain's proximity to Latin America (language), or Israel's distance from the Middle East (cultural similarity).



# THE WORLD ACCORDING TO TIKTOK

## MAPPING DIGITAL BOUNDARIES AND ALGORITHMIC PROXIMITIES

### PROJECT DESCRIPTION 2/2

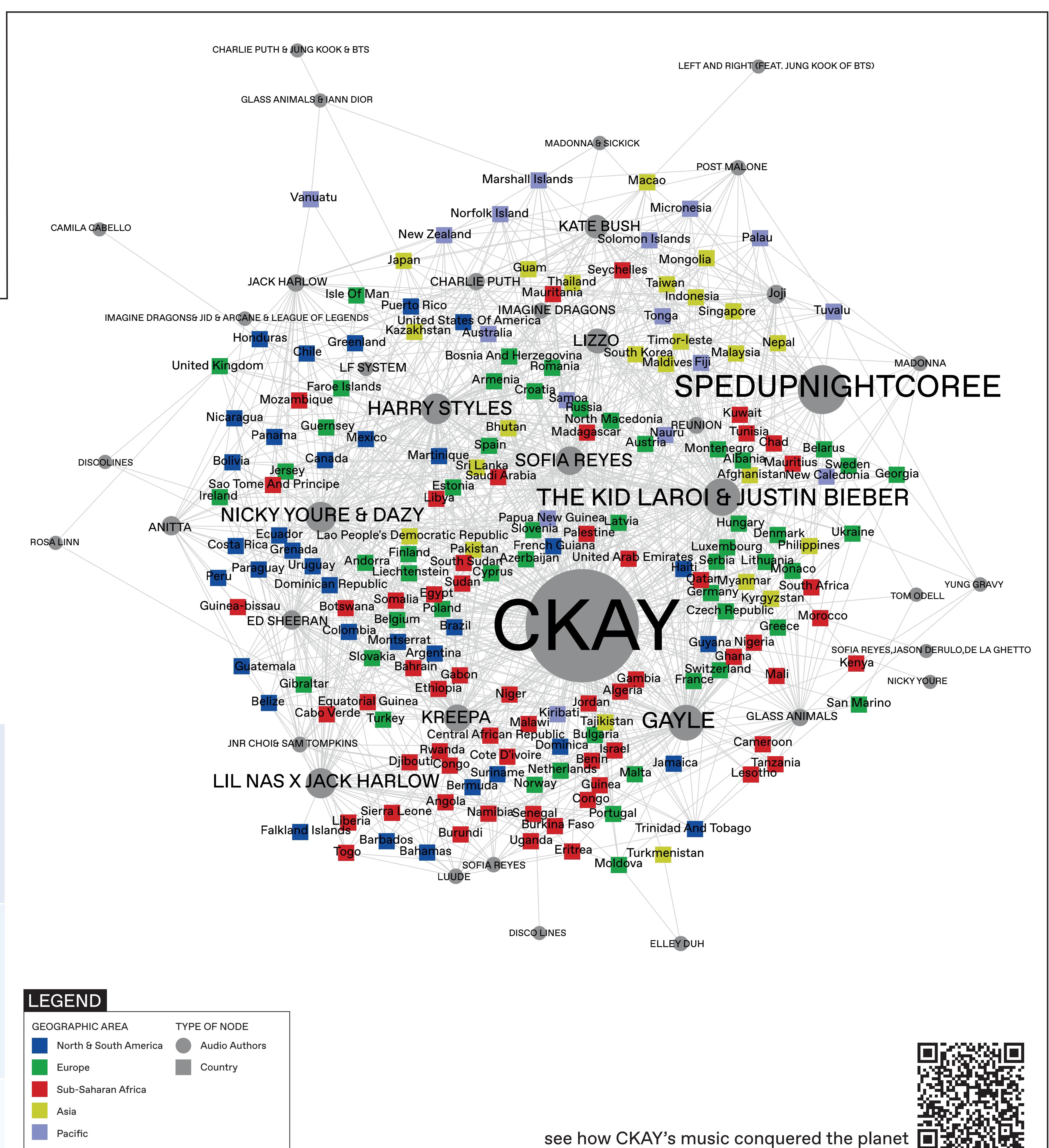
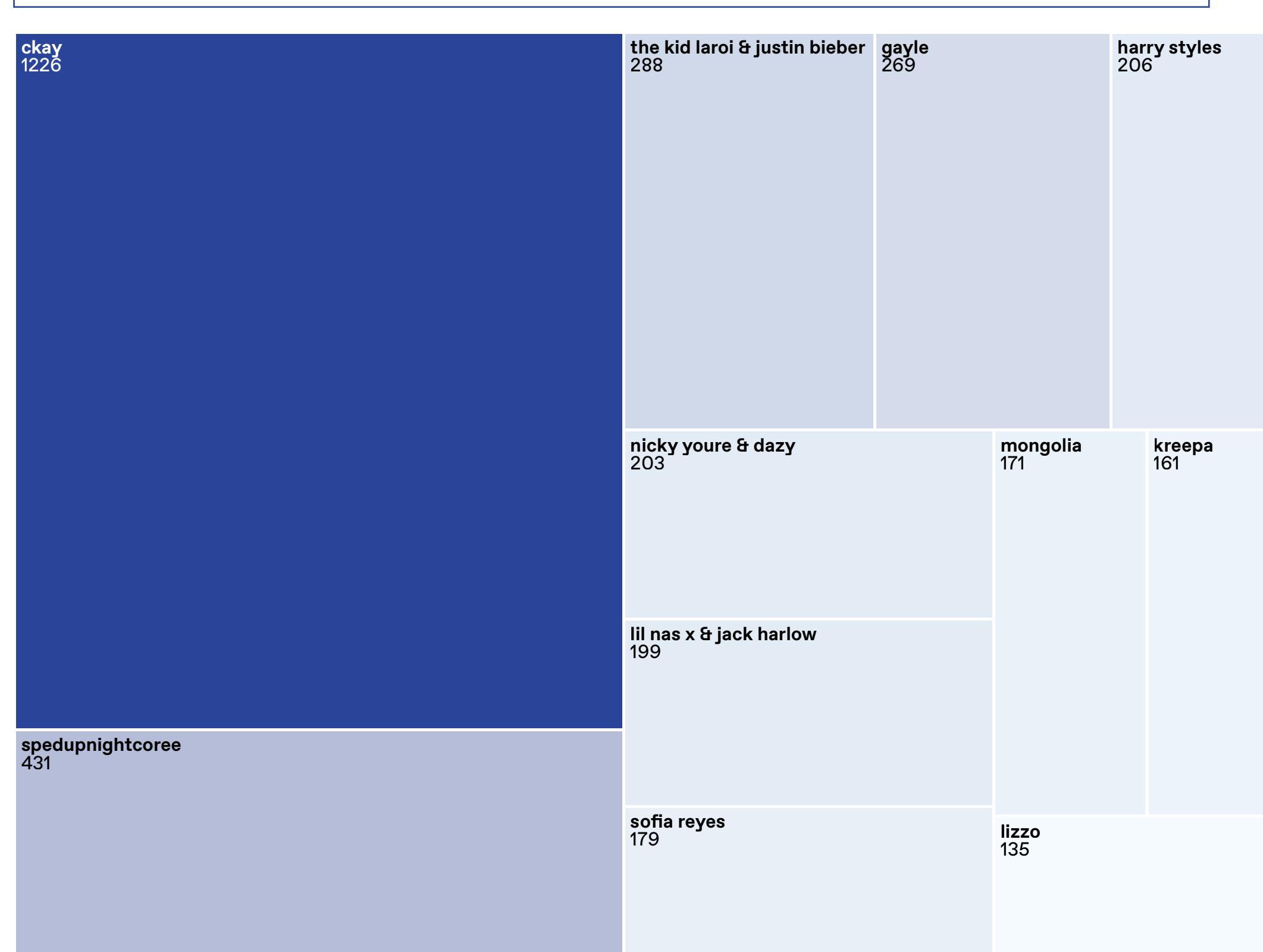
The second half of the project moves away from the macro-level and zooms in to the level of content on the FYP. This is an important step as the content is what the user immediately experiences when opening the app. Two of the variables that we identified are the video descriptions and audio. Focusing on how users describe their own content offers insights into the systems of categorisation that are presented to the user. Additionally, TikTok virality has impacted the music industry. It is therefore important to analyze the mechanics of how the platform promotes music on its FYP. By focusing on the exposure and circulation of audio and labels globally, the following projects visualize the circulation of specific content and are a step towards understanding their sociopolitical implications.

### SOUNDS OF TIKTOK

After an initial Gephi mapping of sound across countries, we found strong homogeneity of sounds: most countries are exposed to the same TikTok sounds. Hence, we proceeded by mapping particular genres onto the various countries. However, most of the 'songs' on TikTok are actually internally produced TikTok 'sounds' that are then reused. Consequently, we decided to focus on artists specifically: we found a list of top artists on TikTok in 2022 and searched for the top 50 top artists in our own data set, of which 30 were present. In Gephi, we then matched the artist to the countries, and lastly, we categorized them per continent, resulting in our final graph. We also made a gif showing the changes over time in the circulation of the most frequently occurring artist in our dataset: CKay.

### FINDING

Our first finding was that most of the videos that were presented consisted of TikTok original sounds rather than music of actual artists, and these were similar globally. Even when the dataset was filtered for just pop-artists, users still experience the FYP in a sonically similar way, being exposed to roughly homogeneous music regardless of their location.

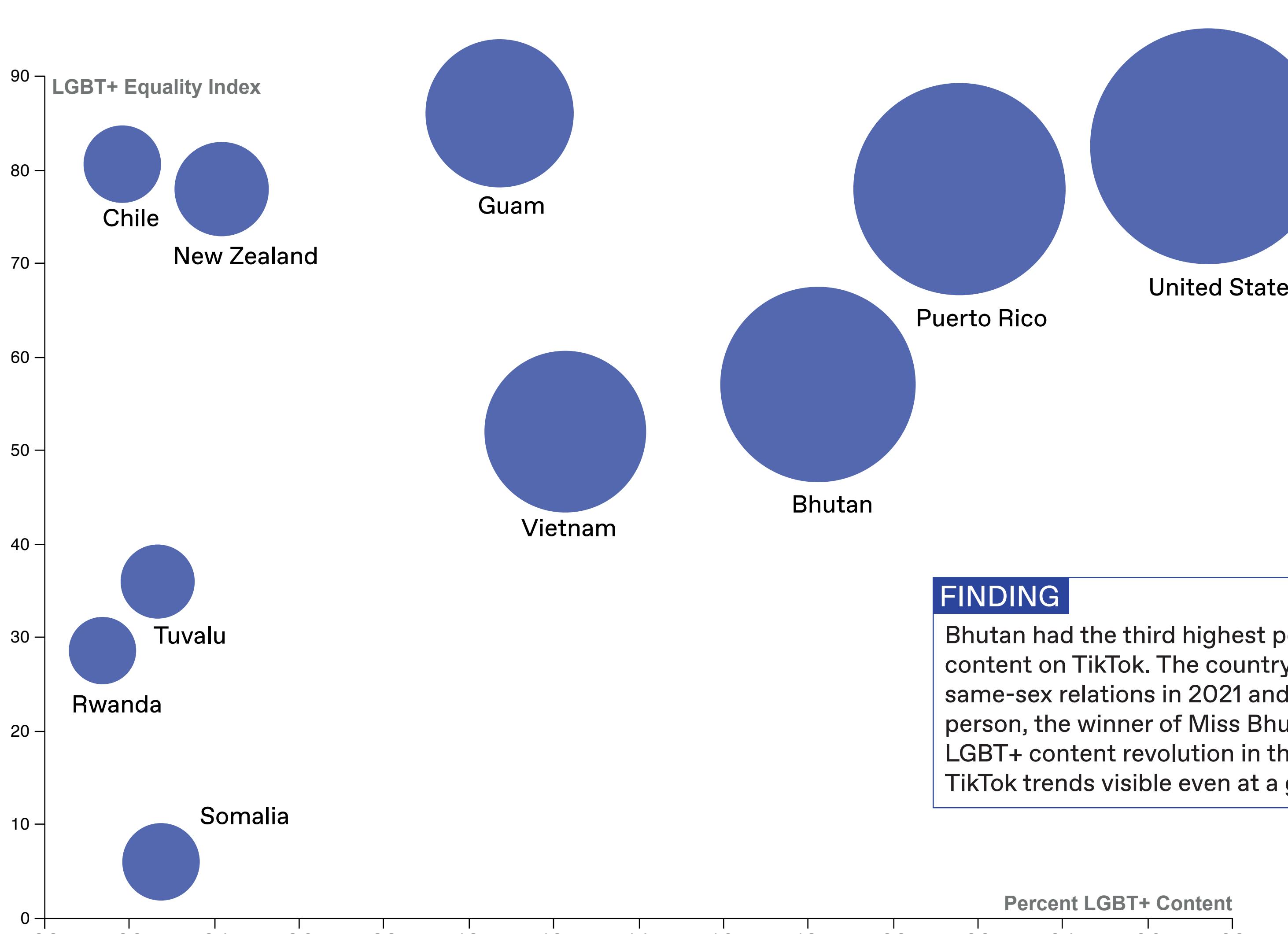


### EXPLORING QUEERTOK

Despite TikTok taking an apolitical stance, the platform algorithmically serves political content. We start by exploring the representation of human rights issues in the videos the TikTok algorithm pushes to geographically diverse FYPs. From there, we sample for videos which specifically contain LGBTQIA+ related words and phrases in the hashtags and descriptions. We find that FYPs in the US, Puerto Rico, and Bhutan showcase at least 1 in 50 videos containing LGBTQIA+ related terminology in the description. Then, we perform a qualitative sentiment analysis of the remaining videos and remove all spam and irrelevant content. We find that all content was positive or neutral, with no homophobic content.

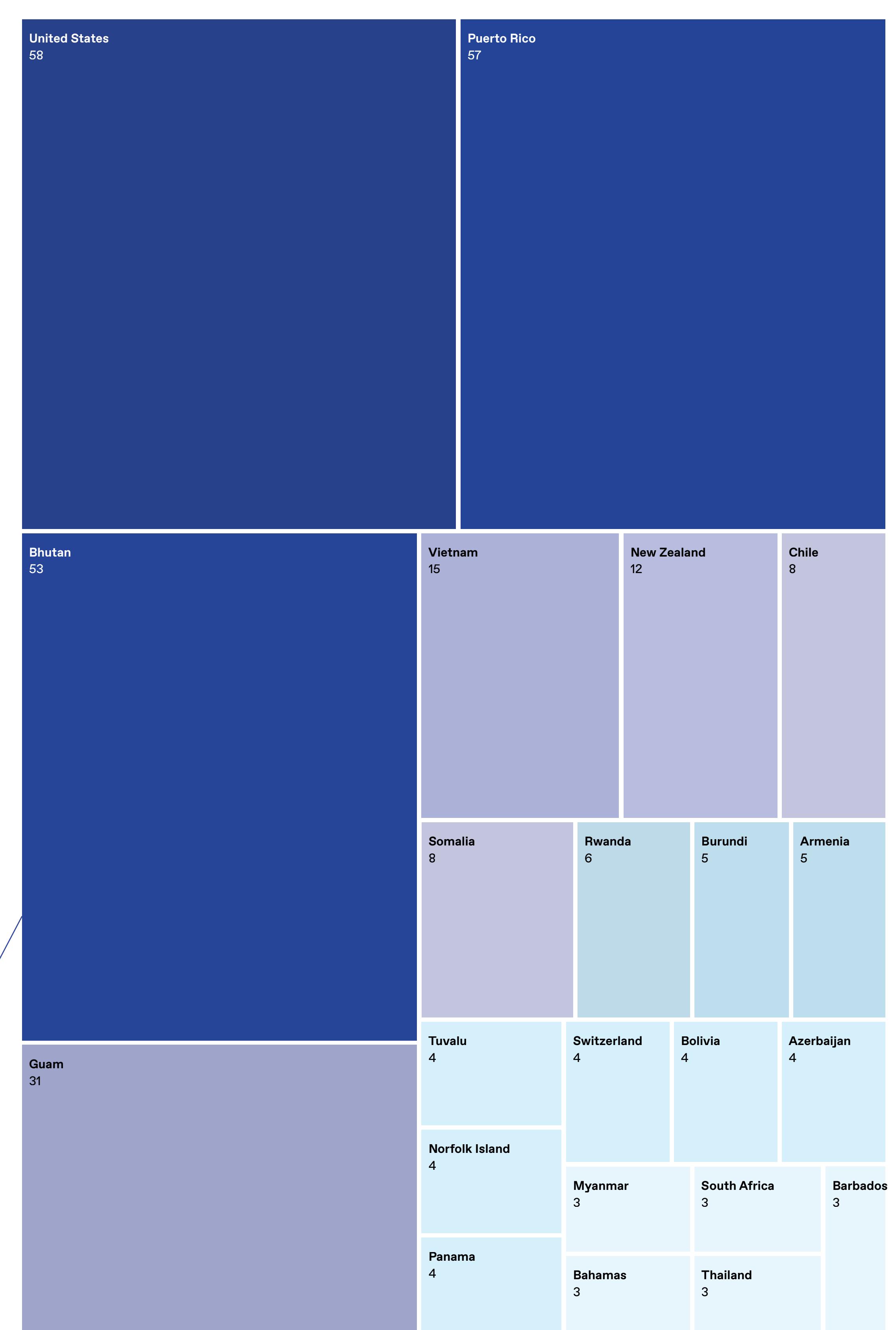
### FINDING

Our analysis reveals that even countries scoring relatively low on EqualDex's LGBT+ rights have a noticeable proportion of queer content circulating.



### FINDING

Bhutan had the third highest percentage of LGBT+ content on TikTok. The country decriminalising same-sex relations in 2021 and an openly queer public person, the winner of Miss Bhutan 2022, sparked a LGBT+ content revolution in the country. This drove TikTok trends visible even at a global level.



# THE WORLD ACCORDING TO TIKTOK


# MAPPING DIGITAL BOUNDARIES AND ALGORITHMIC PROXIMITIES

# HASHTAGS & LABELS

These two graphs are generic representations of content on regional FYPs, categorized by hashtags and labels. Hashtags are user generated whereas labels are generated by TikTok. As such, they provide two interconnected views of TikTok content that can lead to further investigation.

In the hashtag graph, we can clearly see macro-area clusters, which are similar to what we see in our initial, overarching map of TikTok videos. Our next step will be to clarify the relationship between how TikTok's algorithm territorializes content and its potential influence on how users decide to generate content.

Future research on labels should determine how the platform creates them, why they are missing from particular videos, and the relationship between labels, hashtags, and countries.



**watch our video presentation**