**EXECUTIVE SUMMARY**

**An Average Week Across 2020**

* X Chinese public diplomacy accounts will be active across Facebook and Twitter
* They’ll post X times in total. Y on Twitter and Y on Facebook.
* They’ll receive Y Retweets.
* They’ll attract X comments and replies.
* They’ll provoke ZZZ upvotes.

**DO THEY FIND AN AUDIENCE?**

What engagement, interaction and amplification metrics have we collected? Sum that all here please.

**THEIR MESSENGERS**

China uses a number of different kinds of voices on social media to spread its message: state media, embassies, ambassadors, ministry officials, Confucius Institutes, and Consulate Generals.

* How many of each account type is in the list?
* State media are by far the most active: sending XXX messages.
* Ambassadors are the most amplified

**THEIR MESSAGE**

**THEIR FOCUS**

Can we just sum the number of messages sent by accounts listed into the different regions?

Data Quantity

* How many total messages?
* How many per day?
* Split by year?

Languages

Per Account Types

THEMES

Ahmed

INTRODUCTION

This project has been dedicated to understanding the public diplomatic activity of China on Twitter and Facebook. It has sought to understand how China uses these social media platforms, the kinds of messages that it propagates, the audiences that it might be addressing and the languages that it is addressing them in.

Primarily, the project has been data exploratory in nature, seeking to measure the overall scale of the activity, the main themes of that the messages fall into, and broad patterns over time and across language that can be discovered. Second, it attempts to provide information that will allow the reader to make an informed determination about the overall benefits and challenges of continuing and extending this form of work over time. Third, the report discusses the techniques that have been used to conduct the analysis, their measurable performance, and the implications of the strengths and limitations of these methods for the results themselves. These are the principal questions to which the project - and so the report - are devoted.

METHODOLOGY

The project was divided into X phases.

**Stage 1. Data Collection**

The project began with a series of steps to develop a data collection architecture to collect Chinese public diplomatic activity on Twitter and Facebook.

* First, a ‘Master List’ of accounts was developed by BBCM. These were accounts that were understood by BBCM to each represent a Chinese public diplomatic presence online because they were: LIST ACCOUNT TYPES.
* Using Method52 (see below), CASM undertook data collections for these accounts. For Facebook accounts, the CrowdTangle dashboard was used. For Twitter accounts, data was collected from Twitter’s ‘Steam’ Application Programming Interface. XX accounts were included in the Master List from BBCM, and XX were collectible using these methods.
* The Master List was thereafter used as a single point of truth between CASM and BBCM to define the project’s data collection. It was updated on XXXX to reflect changes in appointments to China’s diplomatic staff.
* This data was presented to BBCM on XXX.

**Stage 2. Identifying Languages and Formation of Language teams**

An initial dataset was collected according to the process set out in Stage 1, above. This was passed through a language XXXX algorithm within Method52, which organised each message into the category of language that it was most likely to be.

These language breakdowns were: XXXX XXXX. These were presented to BBCM stakeholders, and four languages were selected to be analysed: English, French, Arabic and Spanish. BBCM convened teams for each of these languages comprising of fluent speakers in them, with each tasked to proceed to the analytical steps laid out in Stage 3.

**Stage 3. Identifying Themes**

The next stage was to undertake a process to identify the themes that Stage 3 would seek to identify. The process to do this was inspired by a well-known sociological method called ‘grounded theory’.[[1]](#footnote-1) This process emphasises the importance of analysing data using categories that are iteratively and constantly drawn from the data itself by analysts working face-to-face with it. In general, this tends to reconcile the the researcher’s idea of what comprises a category with the actual data itself, creating an eventual framework which reflects both.

To begin this process, an initial series of themes was shared by BBCM. Sessions were then held with each of the language teams to appraise 1,000 randomly selected messages. As new messages were seen, they were either placed into an existing category, or caused a new category to be created.

This process lasted for roughly a month, which created an updated themes list. These were then centralised, and presented to BBCM, and a final thematic framework was created that reconciled the individual products of each of the language teams. These were:

* Insert Themes

**Stage 4. Building the Keyword Thematic Annotators**

The project next moved to building keyword annotators to categorise messages into the themes identified in Stage 3. Keyword annotators are analytical components within Method52 that place any message into a category if they contain a given word. To do this, the same process was conducted with each of the language teams:

* 1000 randomly selected messages were sent to each language team, and marked into each of the themes (or indeed none of them) that were identified during Stage 3.
* CASM used a statistical language analytical count called TF-IDF to identify keywords that most correlated with messages marked into each of these themes.
* These candidate keywords were presented by CASM to each of the language teams. The keywords were appraised and either added as keywords or discarded,
* Additional messages were then placed into any of these categorises if they contained a keyword identified through the process above.
* These were apprised by the country teams, and keywords were both added and removed. This iteratively created a series of additional versions of the keyword lists.
* The performance of each version was measured (see below) and the keyword lists with the highest performance were used to produce the final analysis.

**Stage 5. Testing the Keyword Thematic Annotators**

The measurement of the performance of each of the keyword annotators, for each of the languages, was an important part of the project. In order to do this an evaluation dataset was created for each language. The language teams were given an additional set of XXX messages, which they manually categorised into themes. Then, whenever the keyword list changed, it was used to also categorise the messages on the list. The automated and human categorisations were compared, with the automated result being correct if it agreed with the human categorisation.

This measurement of the keyword annotators produced three important metrics:

* *Recall*: The number of correct selections that the annotator made as a proportion of the total correct selections it could have made.
* *Precision*: The number of correct selections the annotator made as a proportion of all the selections it has made.
* *Overall, or ‘F1’*: All annotators are a trade-off between recall and precision. Classifiers with a high recall score tend to be less precise, and vice versa. ‘F1’ equally reconciles precision and recall to create one, overall measurement of performance for the annotator.

The overall performance of all annotators across all themes and languages was XXX. The breakdown of these results is shown in the following table:

|  |  |  |  |
| --- | --- | --- | --- |
| **Language** | **F1-score** | **Precision-score** | **Recall-score** |
| Arabic | 0.519 | 0.626 | 0.482 |
| English | 0.566 | 0.627 | 0.545 |
| Spanish | 0.47 | 0.616 | 0.425 |
| French | 0.476 | 0.434 | 0.573 |

**Stage 5. Reporting and Visualisation**

Finally, the final keyword lists were used to categorise all of the messages that were collected. The results were exported into a third-party data visualisation package, and visualised for this report.

**Research Ethics**

Issues of privacy online are complex. In some cases, online spaces are clearly public - such as Twitter’s timeline- or clearly private - direct messages on Facebook. In some cases, the privacy of some spaces is more ambiguous, as with open groups on Facebook or very large fora where membership is required. In addition to this, public perceptions of what social media spaces are public and which are private can vary significantly from the legal reality or terms of service. Because these discrepancies between reality and perception relate to issues of autonomy, where information that is not public or information that might reasonably be perceived as private is sought within a research project, the acquisition and recording of this data must be well considered, justified and documented.

Given that the phenomenon studied in this report is public diplomacy, we assess that the individuals producing the information have very little reasonable expectation of privacy; their messages are entirely public and indeed intended to be read by a wide audience. However, no activity undertaken by any individual or institution that is not a verified Chinese diplomatic presence was collected for project or is presented in this report.

**Technology - Method52**

The core technology used in this project was called Method52. It has been built to allow people from outside of formal data science backgrounds to collect, analyse and visualise datasets that are very large and unstructured. This is especially the case for large, text-based datasets, such as those drawn from social media, but has also included datasets comprised of emails, forum data, and internal and proprietary data held by large organisations.

The design principle of Method52 is to create development environment through a graphical

user interface. Users select, configure and connect a number of components to create a

bespoke pipeline that data flows through. Each of these pipelines is designed to perform a

particular task and often a number of pipelines are themselves connected together to fulfil a

particular research-driven function. There are 82 components in Method52, and many of these can themselves be configured to perform a number of different tasks.

pasted-image.tiff

**RESULTS**

**Part 1 - CHINA’S PUBLIC DIPLOMACY: SCALE**

**- Overall Volume**

A total of 372 Chinese diplomatic accounts were collected from for this report. 226 on Twitter and 146 on Facebook.

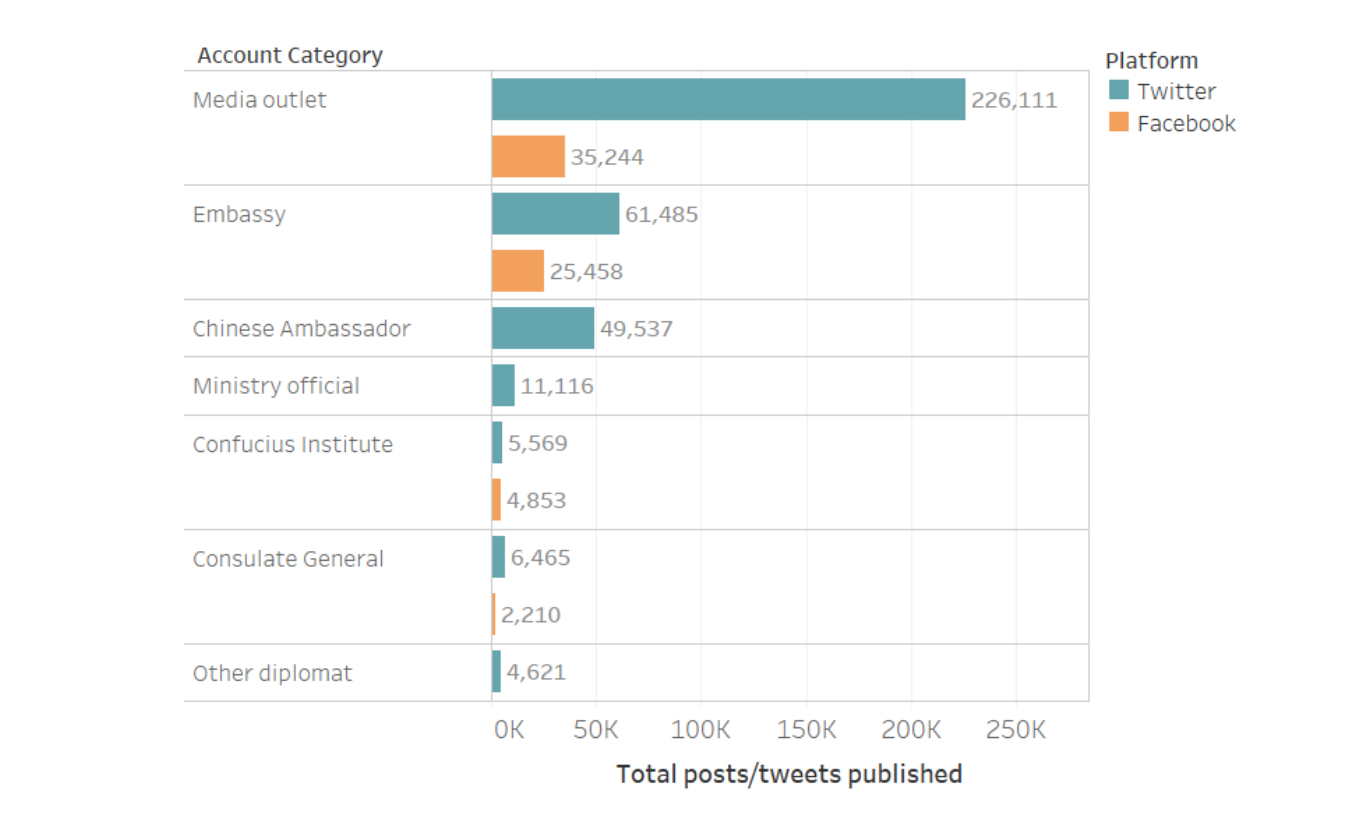
From October 14 to December 8, 2020 (the active period of data collection for this project), They produced a total of 432,800 messages; roughly 365,000 Tweets and 67,800 Facebook posts.

* How many total messages?
* How many per day?
* Split by year?
* Overall volume over time (as VOT?)

**PART TWO CHINA’S PUBLIC DIPLOMATIC MESSENGERS**

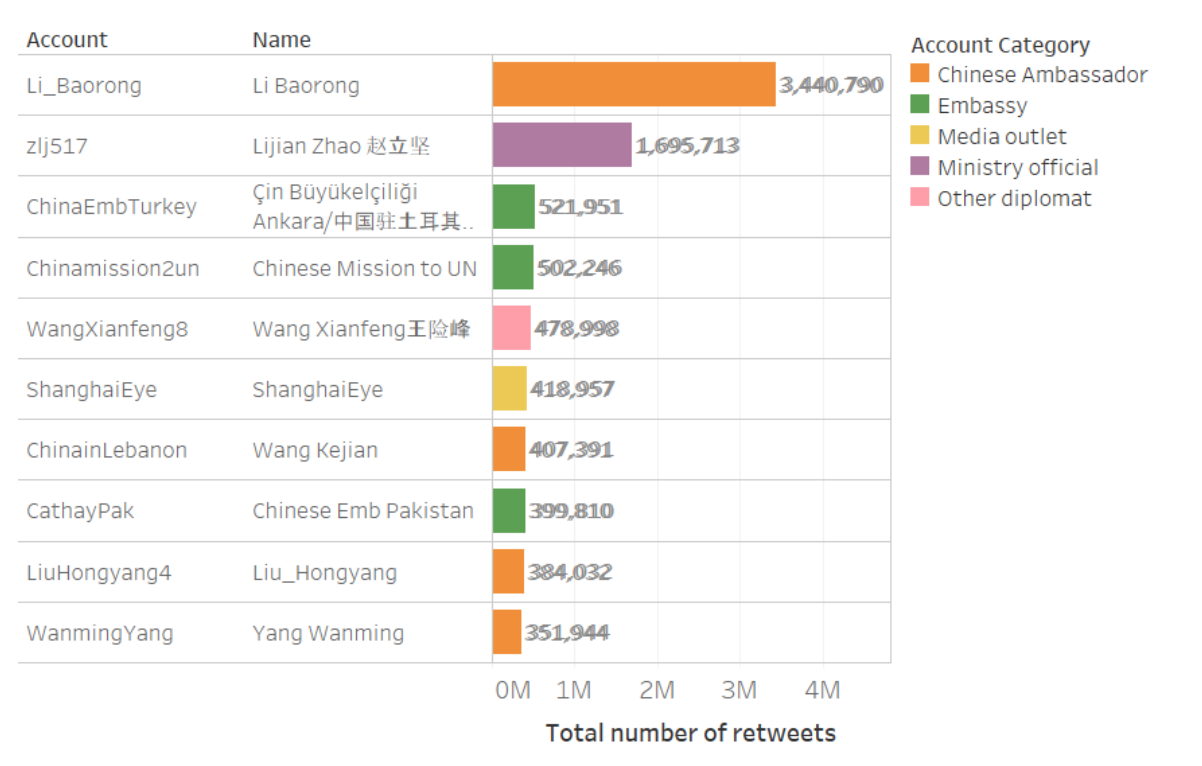
**Account Types**

Each account was placed by BBCM into one of a number of different categories based on the type of public diplomatic presence that it is. State media outlets were, were far, the most prolific kinds of accounts that the project collected from, followed by embassies and ambassadors.



Ahmed

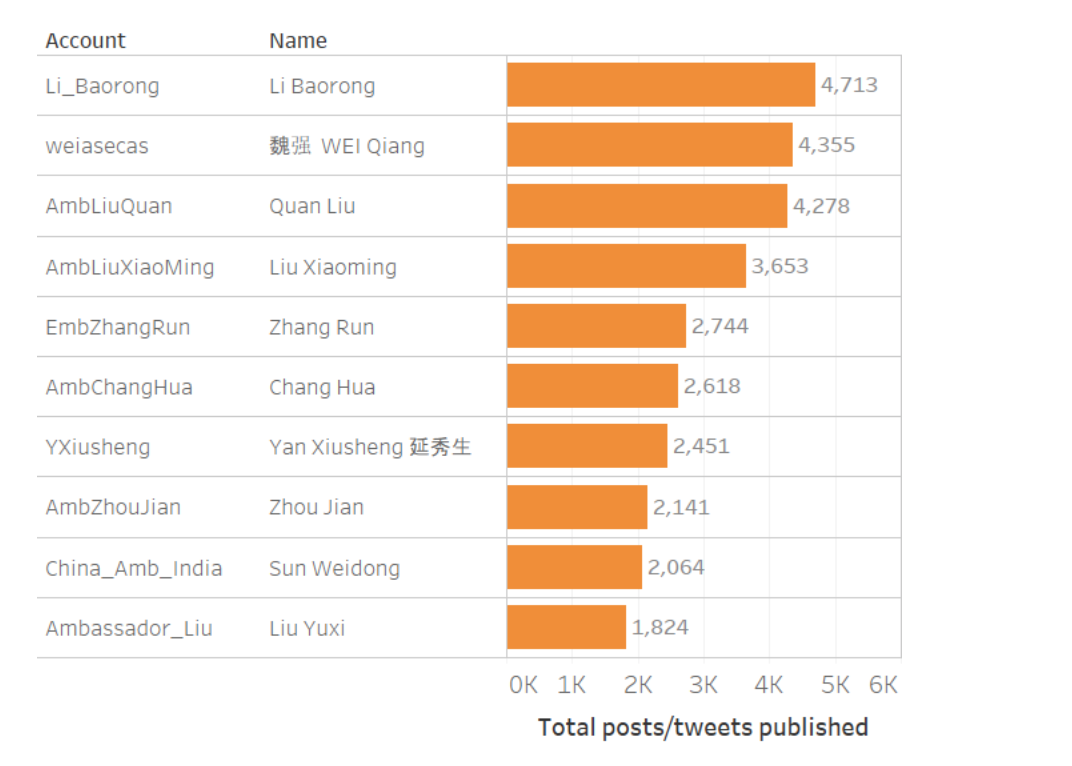
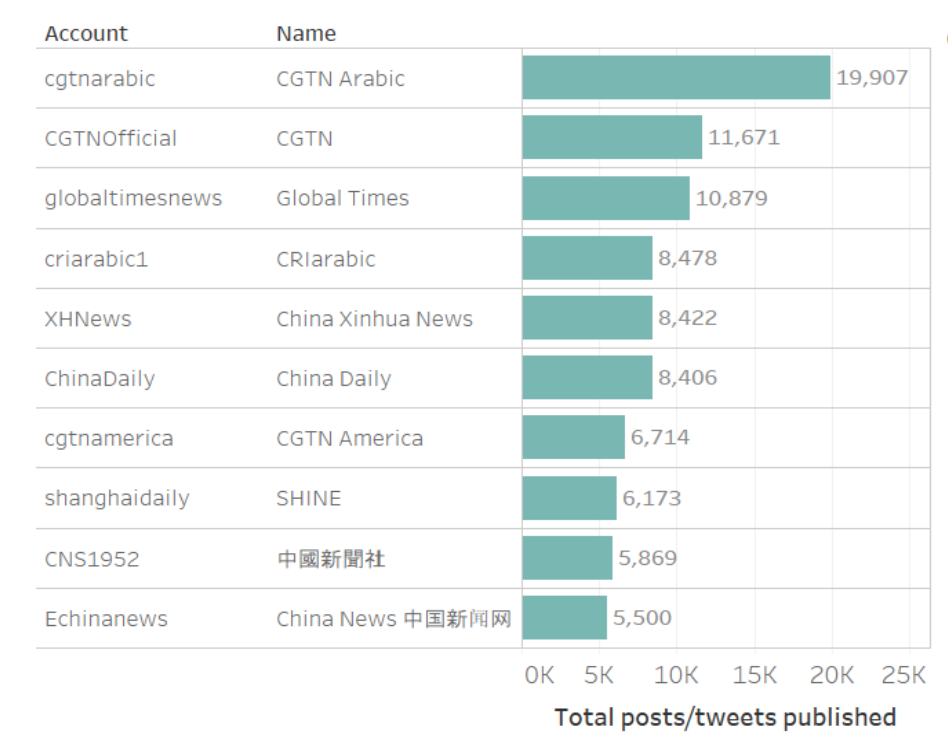
**Engagement**

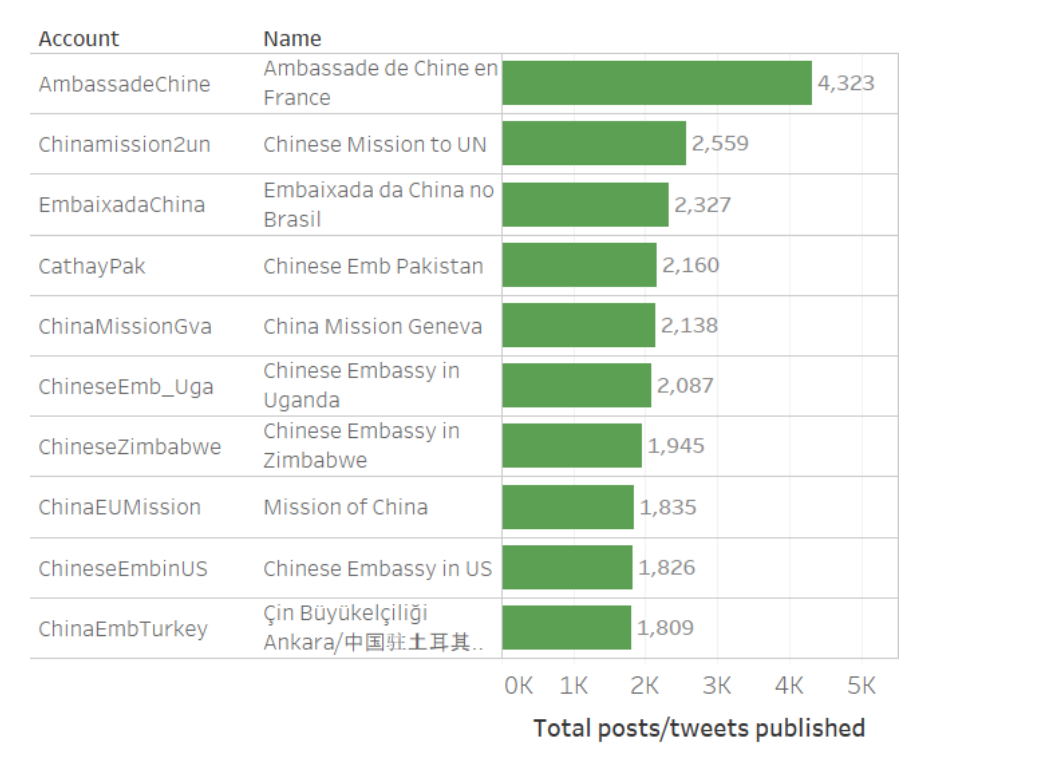
Yet whilst state media generate the most output, they are not the most engaged with. Ambassadors, embassies and ministry officials attract many more Retweets than state media accounts. 

**Prolific Accounts**

Whether state media outlets, ambassadors or embassies, the most prolific diplomatic voices were extremely prolific. Between XXX and YYY The 10 most active across these three categories sent an average of XXX messages, or YYYY a day. This is Z of the total amount collected.

Ahmed





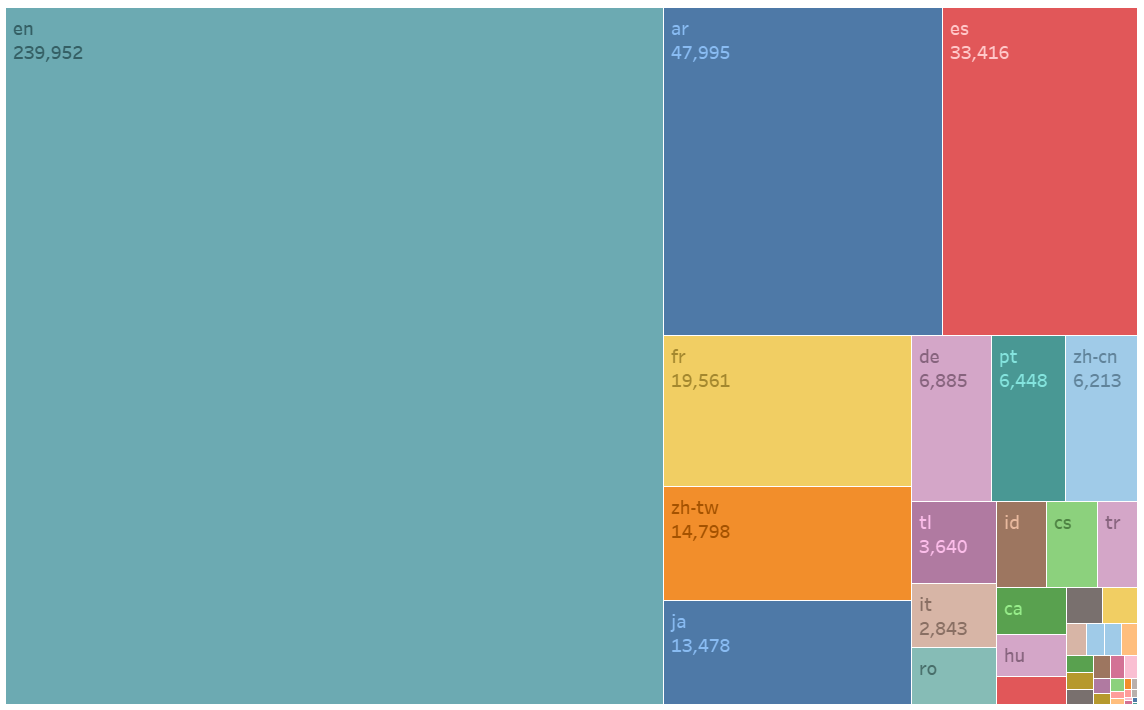
Yet whilst

IS THIS MAINLY STATE MEDIA?

**PART 3 - CHINA’S PUBLIC DIPLOMATIC MESSAGE**

**Languages**

* Most prolific language
* Changes in Languages Over time



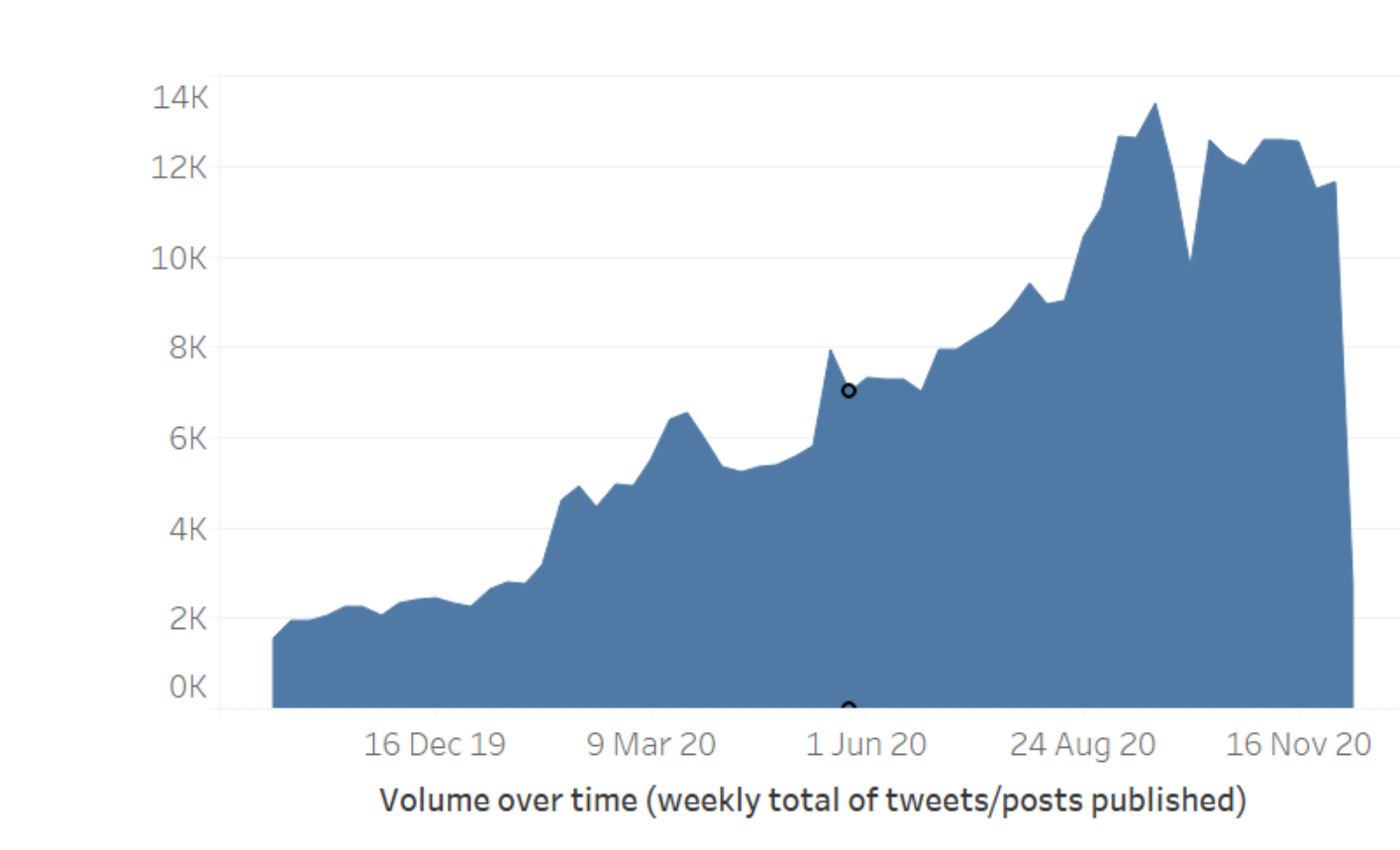
XX

The majority of the messages sent by relevant accounts were in English, but XXX other languages were included too. The most prolific other languages were Arabic, Spanish, French which, together with English comprised the four target languages for this study.

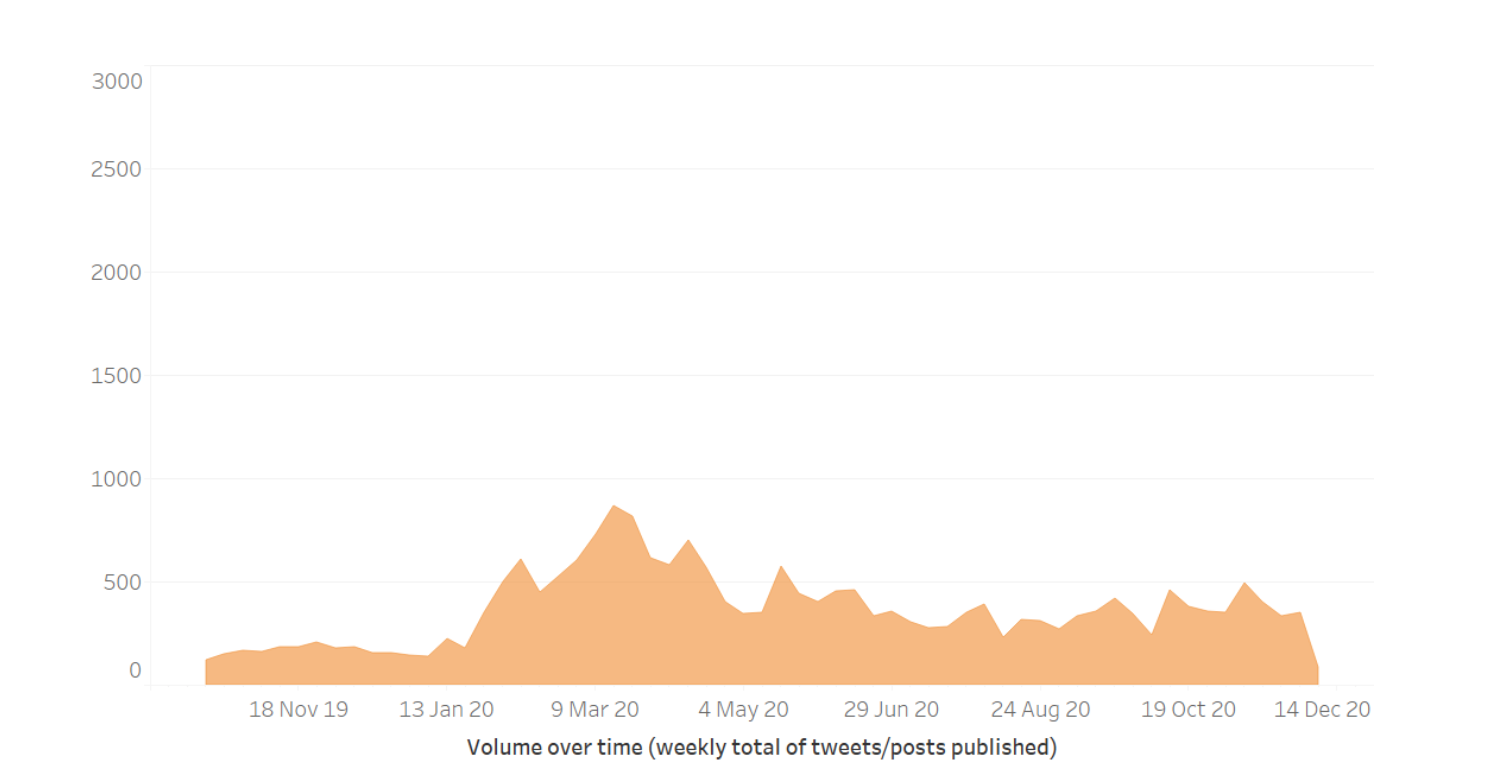
THEMES

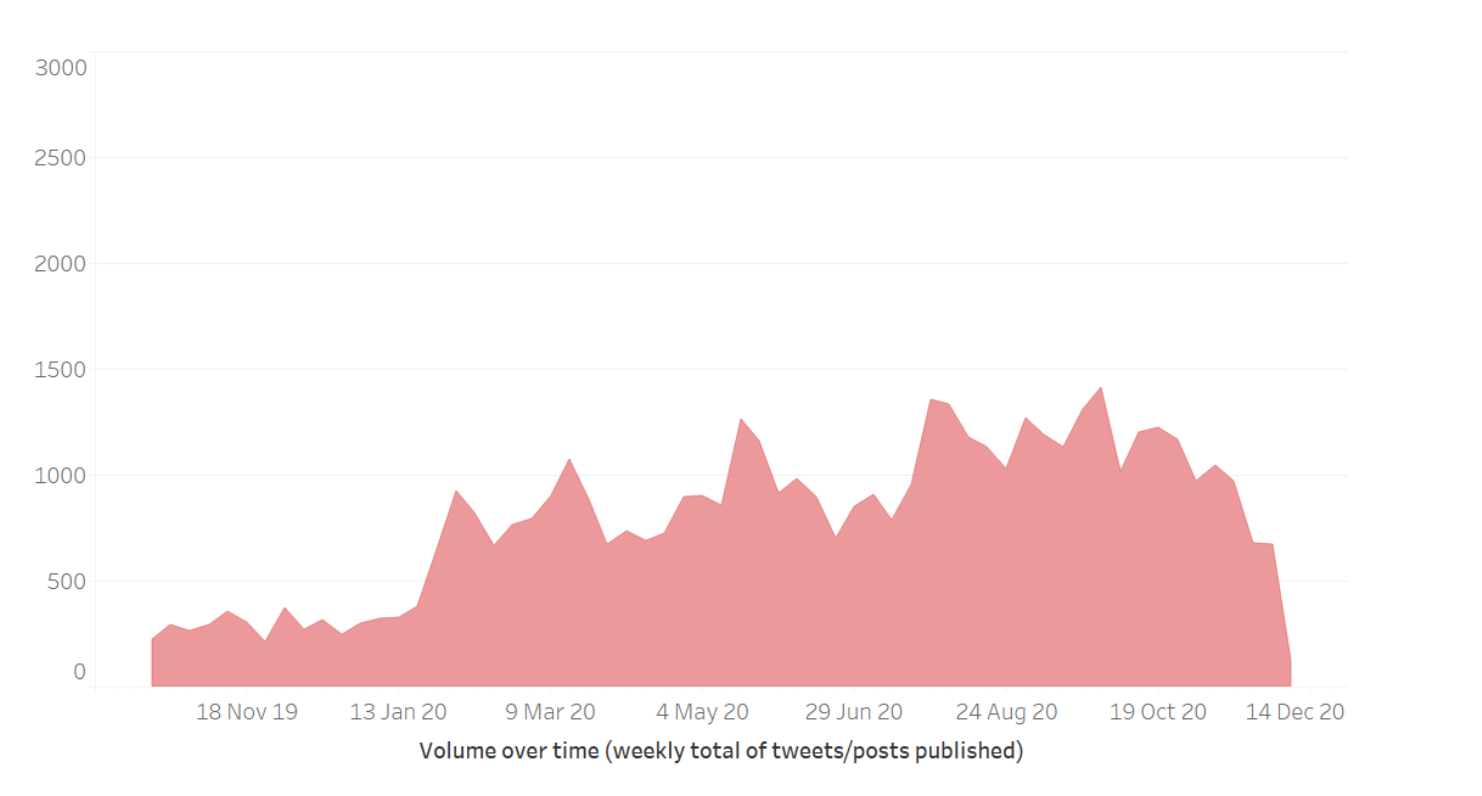
**PART 4 - TRENDS CHINA’S PUBLIC DIPLOMACY’S**

China’s public diplomatic presence has likely been growing since at least the middle of 2019.

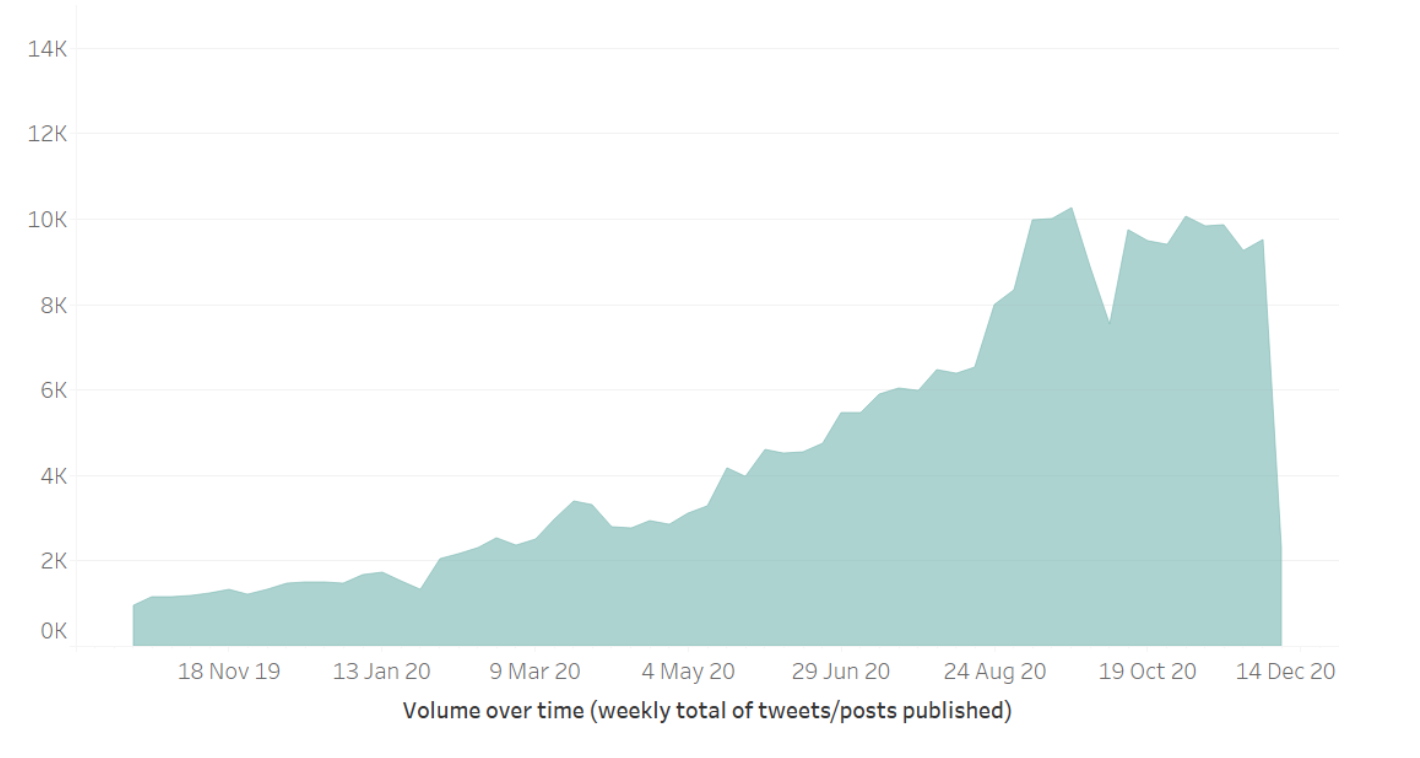


However, the activity of China’s diplomatic accounts based in a number of regions have not increased over this time period. This is the case in Africa.

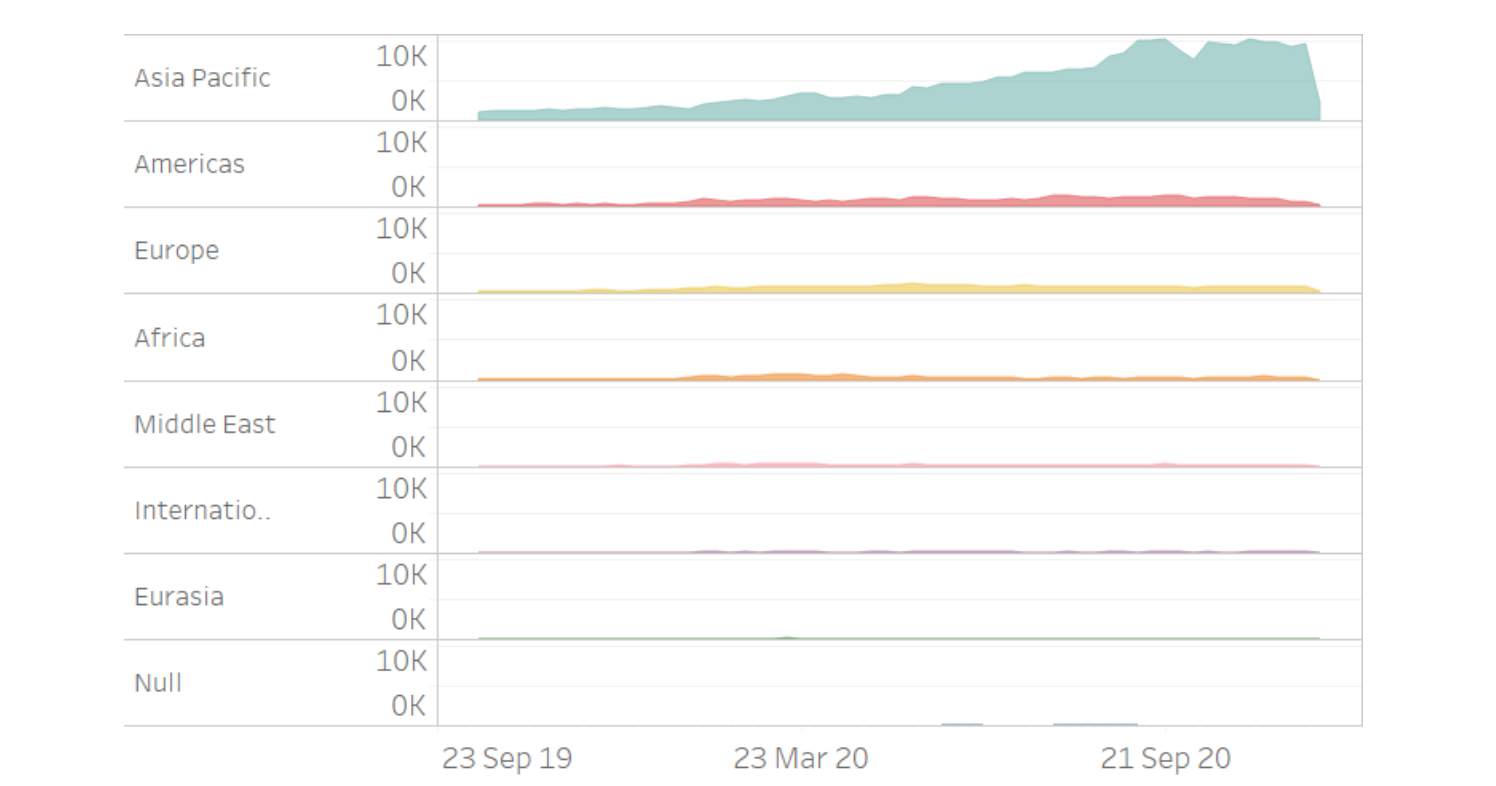


**A reasonably stable scale of activity is also apparent in the Americas**

**It is Chines diplomatic accounts based in the Asia Pacific that has seen both the increase in focus and the most activity overall.**

****

**With each region set again each other, this emphasis on the Asia Pacific is clear**

****

**DISCUSSION - OBSERVATIONS ABOUT THE PROJECT**

The amount of data availability quite high - and Most fell within one of the relevant themes of interest

The performance of the keyword annotators would like be improved through the use of natural language processing

1. [↑](#footnote-ref-1)