Navigating Social Media:

A Case Narrative

December 2021 Primary Technology: Python



Introduction

This report is the first in a series that analyzes the social media environment from several perspectives. This first analysis uses the experiences of Meghan Markle, the Duchess of Sussex as a case study to explore the scale and impact of targeted harassment on social media platforms. Specifically, this analysis examines the sentiment of discussion revolving around a single television event in Sprint 2021. *Oprah with Meghan and Harry: A CBS Primetime Special* was one of the biggest TV moments of the year; that's no small feat in our increasingly individualized media market.

A significant amount of social media chatter can be driven by a very small number of accounts. In October, Bot Sentinel conducted an analysis which found that 70% of negative, hateful, and abusive tweets about Prince Harry and Meghan, the Duke and Duchess of Sussex (D&DoS), were driven by just 83 accounts. Furthermore, through the course of my own analysis, I observed that the Natural Language Processing module used to perform sentiment analysis, like the one for this report fails to capture the complexity and nuance needed to moderate racial discrimination in digital platforms, it frequently misidentified negative content as neutral. Finally, as this project is an independent analysis, there are fundamental weaknesses that represent areas of intended growth. Through the course of completing this project I found elements that I want to finetune, polish, and reevaluate but there were useful insights, so I decided to push ahead with an interesting but imperfect report and return to improve the project over time.

See the appendix for more insights about future modifications to this project and report.

Analysis

Data & Framing

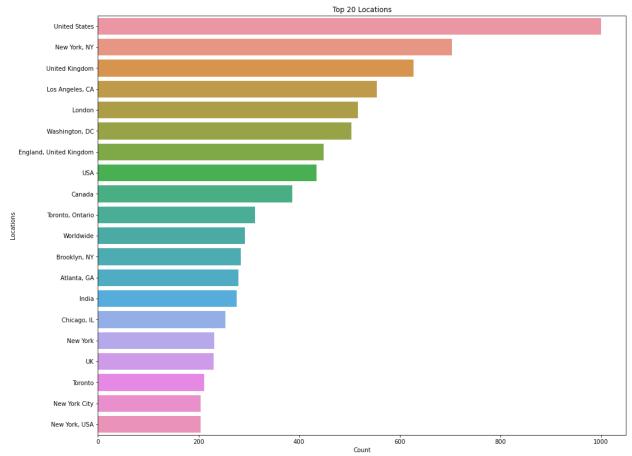
The data for this project was sourced via <u>Kaggle</u>. Content was scraped from Twitter, using the hashtags related to the event to identify relevant data points, surrounding the 'Oprah with Meghan & Harry' broadcast on March 7, 2021. Because the dataset was gathered using hashtags related to the subjects of the special (Oprah, the Duke and Duchess of Sussex, the broader royal family), it is possible that there is a sort of self-select bias at play.

Frequent users, especially intense users organized into fandom groups often called stans, are very practiced using hashtags to hack social media trend rankings. Organizing aggressive campaigns to push their chosen word or phrase to the top of the list. K-pop stans are notorious for taking over a seemingly irrelevant trending topic with gifs, video clips, photos and other content related to their fandom principle or "fave".

The dataset was fairly clean, right off the bat there were 52,802 tweets collected for the dataset. The dataset contained 26,945 unique users or accounts and 8,678 unique locations. The bulk of the code for this analysis was spent parsing out the different elements of a tweet.

After distinguishing "#hashtag" and "@mention" content from among the text of the tweets, I set out to analyze the shape and trends among the data.

Twitter's platform makes location data uniquely challenging. Users are given an open field to describe their location, some of my favorites include "Buckeye Nation", "(D)MV", "nun-ya", and "your mom". The top 20 locations were not at all surprising, high concentrations in the US, Canada, and the UK. Meghan was a lead actress on *Suits* which filmed in Toronto, as such she had a significant Canadian fan base for almost a decade before her marriage. I would like to find a way to group different labels for single locations i.e. grouping America, United States, USA, etc. into one bracket for future analysis.



Follower Ratio and Account Authenticity

As part of this analysis, I hoped to develop a method to detect suspicious accounts. To that end I created a "follower ratio" variable dividing the number of followers by the number of accounts followed. Certain accounts will have high user ratios, these will be major celebrities and publications, accounts that draw a lot of followers but don't have a use to follow others; think major celebrities like Beyoncé who doesn't follow anyone, or the American paper of record, the New York Times. On the other end, accounts that follow more than double the accounts that follow them can be a sign of spam accounts. Botts can be programmed to automatically follow

accounts that meet a certain criteria including tweeting a specific hashtag or retweeting a certain page.

Hashtags

	hashtags	counts
2799	meghanmarkle	33125
3321	oprahmeghanharry	4301
2763	meghanandharry	3366
3615	princeharry	3160
2768	meghanandharryonoprah	3049
1853	harryandmeghanonoprah	2263
3476	piersmorgan	1269
3945	royalfamily	1257
3308	oprah	1233
1849	harryandmeghan	888
2760	meghan	773
3330	oprahwinfrey	494
2141	internationalwomensday	460
1843	harry	454
3634	princessdiana	318
3773	racism	315
2342	katemiddleton	280
648	buckinghampalace	276
1707	gmb	236
3311	oprahharrymeghan	232
	·	

One of the first observation was the frequency of the #meghanmarkle hashtag above all others. Top 10 hashtags cluster between 900 & 4,500 occurances with the exception of Meghan's name, which appeared 33,125 times. Meghan's name was used a hashtag almost 7.5 times more than the next tag which was the name of the special. Of 76,568 different hashtags the top 20 are dominated by direct references to the special, the participants or people mentioned¹ throughout, almost half and 820,224 mentions of Meghan specifically.

Statistically, the Megan Markle hashtag is a significant outlier among the tags related to the televised special. It is not surprising that her name would be part of a conversation as she was a major principle. It was, however, interesting to realize the outsized scale of her presence. The Duke and Duchess of Sussex draw millions of eyes to issues and programs worldwide but, in this case at least, Meghan seems to be the one driving much

of the conversation².

Comparing bar graph visualizations before and after dropping the outlier demonstrates the power of outliers to shape the perception of data³. Dropping Meghan's name for the second figure significantly impacted the shape of the bar-graph. While dropping the #meghanmarkle tag does technically make the data more usable, being able to compare and contrast the two graphs goes a long way to demonstrate the extent of the focus on Meghan.

While the top 20 trending hashtags are not surprising, they do present an interesting insight into the conversations happening around the special. The tight focus on the Duke, Duchess, and royal family broadly makes a lot of sense. The appearance of Piers Morgan and GMB - the British morning show he was co-hosting at the time - and his status as the most frequently mentioned user was both surprising and significant.

¹ Piers Morgan is a notable exception. He was not mentioned during the special but he appeared more frequently than any other individual besides the Duke, Duchess, or Oprah. He even appeared more often than the royal family. The morning show he co-hosted at the time of airing, Good Morning Britain (GMB) also appeared in the top 20.

² Anecdotally, even in instances where Harry and other members of the royal family are the topic, Meghan looms large. The Duchess seems very aware of this fact and often points that power in the direction of charities and causes she supports.

³ See figures 1 and 2 in the appendix

Mentions

Of the top 20 mentioned accounts, 7 have a direct connection to Piers Morgan. At the time he and Susanna Reid (@susannareid100) were co-host of GMB which airs on ITV. Alex Beresford

mentioned_users	total count
@piersmorgan	2032
@Oprah	1241
@GMB	551
@RoyalFamily	375
@Bethenny	202
@KensingtonRoyal	181
@ITV	175
@CBS	161
@alexberesfordTV	154
@DailyMailUK	126
@MailOnline	125
@chrisshipitv	121
@CNN	120
@CBSThisMorning	118
@megynkelly	116
@scobie	106
@halfadams	102
@SkyNews	97
@BBCNews	90
@susannareid100	87

was the weather presenter on GMB and it was a heated exchange with Beresford about Meghan that led Morgan to walk off set and quit the show. Morgan is also a frequent columnist for the Daily Mail and its online division the Mail Online. He was the most frequently mentioned user, appearing 2,032 times in total. The Duke and Duchess do not currently have any social media presence, so it's impossible to say if his shadow would extend so far if the principles of the special were also online.

Chris Ship, Megyn Kelly, the Daily Mail/Mail Online are all part of the media landscape that has been deeply critical of the Duke and Duchess of Sussex, especially their stance on media intrusion as relates to public figures and celebrities⁴. Ship - a member of the royal rota for ITV (which aired the special in the UK) - represents a conventional if critical perspective of the couple; the Daily Mail and the Mail Online are in an ongoing legal battle, appealing their loss against the Duchess over a private letter they published after buying it from her father; and Kelly, a conservitive commentator formerly of Fox News, has occasionally chimed in to critique the Sussex's in a vein that is very common in that media ecosystem. Both Kelly and Bethenny

Frankel (@Bethenny) went viral for negative reactions to the special that were considered in poor taste given the discussion of mental health and in the case of Kelley, hypocritical considering she left a hostile work environment not long before the special aired. Frankel and Kelly are the only two of the top 20 users that are more than 1 or 2 degrees of separation from the Duke and Duchess.

Patrick Adams (@halfadams) is Meghan's friend and former Suit's co-star. He was among a handful of Meghan's longtime friends who took the opportunity to speak out in her defence. We will probably never know if the supportive comments were directly approved by the Duchess or intentionally orchestrated, but a wave of deeply personal statements and appearances across social media and morning talk shows appeared immediately after the special aired. Meghan discussed the extent of coordination that happens inside palace walls regarding public communications. The Duchess even said she was reprimanded after her friends gave comments to People magazine in 2019 in an effort to defend and humanize her in

⁴ Many people including media personalities & commentators insist that intrusion is the price they pay for fame. This argument is a major component in the narrative that Meghan in particular is an attention seeking hypocrite; that she court's attention but also "wants privacy".

the public eye. The sudden appearance of such intimate statements in her defence at the same time the couple broke their own silence was an indication that the Duke and Duchess were serious about their departure and was taken as a point-of-no-return by detractors and fans alike⁵.

Omid Scobie (@scobie) is the royal correspondent for ABC/Good Morning America and royal editor for Harper's Bazaar. He is also the author of *Finding Freedom*, the biography of Harry and Meghan that set off controversy. The book remains one of the most high profile positive perspectives on their relationship and events leading up to their departure from the UK. His status as one of the few consistently positive voices in the royal rota led some to speculate that he worked directly with the couple to source the book. This claim would become a central component of the Mail's failed argument against the Duchess in court.

Based on my own observations, I would conclude that the abundance of news channels represented in the mentioned accounts is a result of many users replying to threads of live updates media accounts often create for high profile broadcasts. Direct replies to a tweet connect to the engagement metrics of every account mentioned in the original tweet, unless the respondent actively removes any of the accounts mentioned in the thread. In essence, some of the text captured in this data happened in response to headlines tweeted by legacy media organizations - BBC, ITV, GMB, and Sky News in the UK and CNN, CBS & CBS This Morning⁶ in the US.

Piers Morgan stands out both for the frequency of his references and his place in the narrative. He is often characterized as Meghan's primary adversary in the British media. Morgan has a well documented history of belittling and dismising high profile women's disclosure of their own mental health struggles, claiming that both Madona and Lady Gaga lied about past sexual assault. He accused Gaga of useing her candor about experiencing PTSD to promote her music, and claimed not to believe that Meghan suffered suicidal ideation while pregnant with her son Archie.

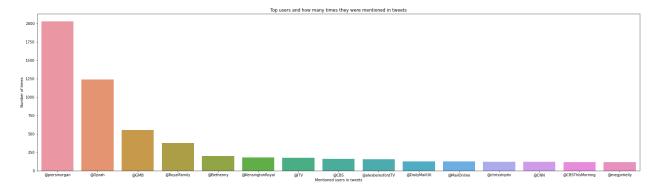
He briefly interacted with Meghan Markle during early periods of her career, used the interaction to inflate his relative importance, and then portrayed her as a <u>fake</u> social climber when she didn't respond to his later provocations. Over the course of Meghan's courtship, engagement, and early marriage, Morgan <u>distorted</u> a 2016 exchange during the Suits press tour in London, claimed to have put Meghan into the cab that drove her to a blind date with Prince Harry, used his media platform to ask for an invite to the wedding, the wedding takes place and Morgan immediately begins criticizing her humble "performance" preceding the "scripted"

⁵ It is noteworthy that the palace outlined a 1-year review period, offering the couple an off ramp if they found life outside the palace less than ideal. The fact that the special was announced almost immediately after that review period ended was seen as an indication that when the couple said they wanted out, they meant it the first time.

⁶ CBS This Morning is co-hosted by Gayle King, a long time friend of Oprah. Oprah, who has a long standing relationship with CBS beyond her friendship with King, shared exclusive previews on that morning show before the special aired and appeared the morning after to discuss the.

"hollywood" wedding, and finally claimed to have been jilted *because* he criticized her. The Duchess's silence⁷ leading up to the wedding seemed to enrage Morgan, as after months of trying to get her attention with flattery, he made a hard pivot to a narrative of having been ghosted in favor of Harry almost immediately after the wedding⁸. Morgan further inserted himself into her personal life, frequently inviting her estranged father onto his morning show where the host asked leading questions about how "badly" she behaved and embarrassed the Queen.

Morgan is a prolific twitter user with a decent following, it is possible that some of the engagement resulted from people retweeting and replying to his commentary regarding the special, but that cannot be confirmed without further analysis. Piers overshadowed anyone directly mentioned, associated, or involved with the special, perhaps suggesting that users were bringing him into the conversation. Supporters of the Duchess, the most ardent among them known as the SussexSquad, avoid using his name or twitter handle because they have a demonstrated capacity to make people and subjects they discuss trend on Twitter. Again, fan behavior is incredibly sophisticated, the squad find many colorful euphemisms for people and institutions they consider against Meghan specifically to avoid contributing to the perceived relevance of her detractors. Regardless of how he became such a prominent voice in the chorus, the "mentions" tables suggest that Morgan has managed to be part of Meghan's story, whether she wanted him there or not.



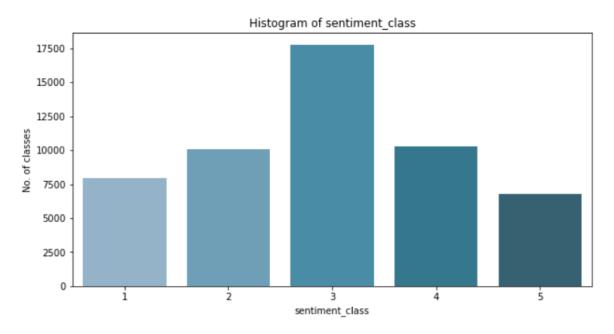
Sentiment Analysis - Natural Language Processing

All of the preceding analysis is designed in service of a larger sentiment analysis. As a first attempt, this project meets its intended goal. I was able to perform a sentiment analysis on the text of the tweets gathered into the dataset. The sentiment analysis was performed using the nltk, gensim, and sklearn packages. After performing the analysis (see the attached jupyter notebooks for the relevant code) the modules concluded that the text was largely neutral, with a normal distribution around that median value. The sentiment classes are organized as follows:

⁷ Meghan would go on to tell Oprah that even in instances where she wanted to interact with or respond to the media, she was told not to as soon as her relationship with Prince Harry became serious and she became a part of his comms strategy.

⁸ He explicitly calls her a social climber, implying if it weren't for Harry's status, he would still have access to the woman he once called "the perfect princess".

1 & 2 = negative, 3 = neutral, 4 & 5 = positive. The sentiment classes are almost perfectly symmetrical, on the ends, class # 1 - the most negative content - slightly edges out its mirror opposite, the most positive content.



Limitations, Planned Updates and Improvements

This stage of the project is designed to focus on the code, the process of getting to the analysis. While assembling this report, I realized there are a number of tasks I want to perform that would be useful or relevant. For example, a table with the value counts by sentiment class would be more effective than a graph to determine how many tweets were in each class. Additionally, the text of the bar charts are very small. It later occurred to me that flipping the "hashtag" and "mentions" frequency visualizations to match the "locations" graph might improve readability. If that is not enough, the update will limit the number of values on the x-axis so the labels have more space. I also want to run this project through Tableau which makes interactive visualizations simple to produce. Interactive visualizations would allow the bar graphs to provide additional details when the user hovers over the metric they want to evaluate. Likewise, the tables are visually inconsistent and the updated report will consistently format them.

Additionally, because the dataset used hashtags as the primary data source, it's possible that a more general keyword search would yield a different analysis. Hashtags are a very effective tool for manipulating trending algorithms and stans are well practiced at using them for that purpose. They use copious hashtags to push their faves to top of trending lists. K-pop stans, Nicki Minaj fans (Barbz) and Beyoncé fans (the Beyhive) are particularly effective at using this strategy. The average user might notice their favorite artist trending and think hundreds of

other people are discussing new music or an appearance, only to realize that dedicated fan pages have organized their users to push relevant tags up the leaderboard. There has been less of this behavior of late, it's possible the platform has taken steps to mitigate this particular form of manipulation, but hashtags might represent a tool of self-selected, highly engaged users. I would like to explore this possibility in the future.

The goal in creating the follow-ratio is to combine this ratio with other metrics that can be indications of spam, bot, or other inauthentic activity to create a "suspicious account" category. The next step would be to use the category of suspicious accounts and see if they demonstrate any noticeable behavioral patterns throughout the special. Unfortunately, I have not yet scripted that full attribute, so that section of the analysis will come later.

Finally, the project uses the VADER library to analyze the tone of the tweets in the dataset. I am not sure what the industry standard tools are at social media platforms, but some of the examples of sentiment classification from a random sample raise questions about the capacity of such libraries to effectively moderate human language. During a podcast segment called "Good use of polling or Bad use of Polling" Nate Silver of the 538 Politics Podcast discussed the unreliability of this type of analysis. The complex nuances of human language make this type of analysis particularly difficult if the lexicon is not precisely calibrated.

Conclusion

I was surprised by the amount of work necessary to make tweets iterable and useful in python. For that reason, this analysis does not go quite as in-depth as I hoped, but is an introduction to the tools and practices involved with such projects. The broad conclusion is that the discussion around the CBS special was generally neutral. When graphed, the tweets have a normal distribution around the median sentiment value. However, much of the analytical set up, including identifying mentioned accounts, is in anticipation of more sophisticated sentiment analysis down the line. As indicated in the planned improvements, I would like to complete the suspicious account metric and see if there is any difference between how authentic and potentially spam accounts discussed the special.

A random sample of tweets organized by sentiment grouping allowed me to understand how the VADER library analyses and categorizes the text. One of the tweets categorized as neutral calls Prince Harry "Harry Hewett" - a reference to a longstanding claim that James Hewett, a red-headed man his mother was rumored to have had an affair with before her death, was his father. The tweet in question goes on to claim that the Duke and Duchess should have their titles removed. Experts on the monarchy and the British constitution have said on a number of occasions that the removal of titles in the current context is highly unlikely and unprecedented; nevertheless, it remains a common demand of twitter users who often oppose Harry and Meghan. While the word choice is not particularly negative or hostile, contextually the tweet exemplifies an extremely negative opinion about the Duke and Duchess, including a belief they are essentially illegitimate and should be treated as such.

Likewise, a tweet in another sample says in part: "#MeghanMarkle is salty AF cause her #moonbump 'Archie' didn't get a (racist) royal title" is categorized as neutral. The reference to "moonbump" is part of the persistent conspiracy theory that Meghan was never pregnant and did not carry her children. It is a dangerous conspiracy that not only dehumanizes Meghan, but her children too. The "moonbump" label often accompanies claims that her children are dolls, "borrowed" or "stolen" from other mothers. More concerning still, in the past this theory has been accompanied by a stated desire for someone to attack Meghan while she was pregnant, presumably unmasking her "deception", proving to the world that she is a liar. That context is not neutral, it's dangerous.



Intensity meets magnification—the moonbump/death-threat nexus

The moonbump is a perfect example of the difficult task of social media moderation. The word moonbump has no negative connotations, but the situational context is deeply hostile to the Duchess. The fact that it was included in a tweet that the library defined as neutral demonstrates the need for a more contextual understanding of language, especially harassment and hate speech. If the VADER lexicon is consistent with industry tools, it comes as no surprise that women, women of color, and black women in particular, have such a terrible experience with platform moderation; the tools currently available are insufficient to meet this moment.

Central to this issue is the meme-ification of online trolling and hate. Antagonistic users can use images, euphemisms, and other indirect and non-linguistic methods to harass other users that would not be captured by sentiment analysis or other AI unless they are programmed with the requisite knowledge. Furthermore, the first example from the random sample of sentiment classifications is what I would call a deep-cut. Only someone who pays enough attention to the

⁹ I had taken a screen grab of the tweet in question, but have not been able to find the particular example and the original thread has been removed from the platform.

drama of the royal family would know that the "Harry Hewett" dig is a call back to the 1980's and 90's during the height of "The War of the Wales'", when Princess Diana's spats with Prince Charles played out across the british tabloids. Her rumored affair with a man named James Hewett has followed Harry for decades, but only in the context of people so deeply devoted to their disdain for him, that they believe they can make a case for why he is not a legitimate royal. None of the words in that tweet are particularly offensive, but the context is highly significant. For both analytical and moderation purposes, a static library would not be able to effectively categorize such examples as each community environment develops their own history, language, and lexecon of disparaging contexts over time.

Why does it matter?

The 2020 documentary <u>The Social Dilemma</u> from the Center for Humane Technology delves into the technology that drives social media. Evidence has piled up that social media algorithms prioritize engagements, and our primitive brain is driven to engage with content that makes us angry. Combined with the power of trending algorithms to amplify a small number of voices into a chorus, we have a responsibility to understand how these platforms work.

One of the most revealing developments over the course of this project was the power of a small number of accounts to drive a conversation online. The Bot Sentinel project that pin-pointed 83 accounts driving around 70% of the targeted harassment toward the Duke and Duchess of Sussex, also found that those 83 accounts had a massive reach, they estimated almost 17 million engagements. Their report included observations that one of the most active hate accounts seemed to tweet almost 24/7, an inhuman feat.

Many people understand how to manipulate social platforms in order to draw attention. Some, like pop-culture stans, use their understanding to make their principle trend. Others, like Morgan have learned to weaponize their social media accounts to manufacture a sense of relevance. Worse still, Bot Sentinel has uncovered patterns of coordination, and some evidence that professional trolls are buying fake accounts to bolster their reach.

Several weeks after Bot Sentinel allerted platforms (Twitter and Youtube) to the coordination their system had flagged, their CEO observed a significant increase in <u>positive sentiments</u> toward the Duchess of Sussex. Setting aside the very real issue of people's public and professional reputations. The Bot Sentinel project effectively demonstrated the potential for online perceptions to be shaped by a small number of bad actors. From high profile figures to high school students, the ability to shape perceptions of people and issues, and artificially inflate niche arguments into the scale of world-wide trends, can have real world implications.

<u>Appendix</u>

Relevant Social Media Accounts:

@theroyalfamily - twitter handle for Buckingham Palace, the institution for Queen Elizabeth

@sussexroyal - the Duke and Duchess of Sussex's defunct instagram handle; before the final deal was announced by the palace, it was assumed they would continue to operate the account.

@clarencehouse - twitter handle for Clarence House, the institution for Prince Charles, Prince Harry's father

@KensingtonPalace - the twitter handle for Kensington Palace, the institution for Prince William, Prince Harry's older brother

Top hashtags:

Figure 1-Top 20 hashtags The updated version of this report will rework the scale for legibility.

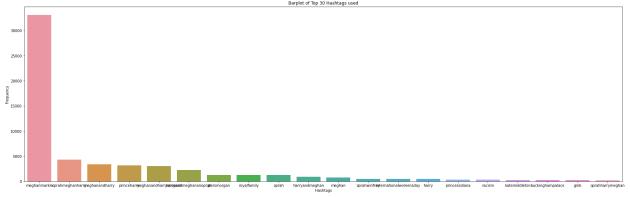
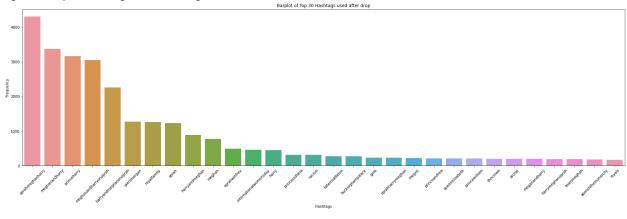


Figure 2 - Top 30 hashtags without #meghanmarkle



[&]quot;Moonbump" example tweets:



According2Taz @according2 taz · Mar 11, 2019

Sorry but it is BIOLOGICALLY impossible to bend down with a baby bump that big, in stilettos, completely unaided and knees together. Before any MEGBOTS slate me, prove otherwise. Get a video of someone as heavily pregnant as Meghan in 6 inch heels to prove it!! #moonbump #megxit



VIDEO: Meghan and Harry depart Canada House after meeting young Canadians on Commonwealth Day #CommonwealthDay



0:27 15.4K views

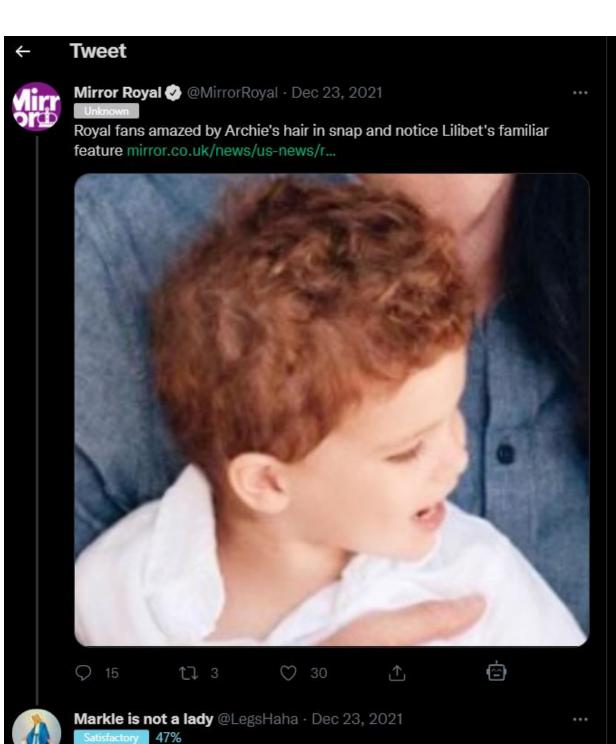


17 19











Hmmmm a borrowed child? Do you not find it odd that child never looks in the camera???? never a photo of him face on?? I smell a rat 🚵

Q 1

17

0 1

 \triangle

