Usage of Twitter vs X to refer to the same-named platform

Joris van Bruggen
S5723752

Abstract

The re-branding of Twitter to X is a rare case of a major company changing the customer-facing branding of a major social media site. To study to what degree such a re-branding actually can succeed in changing the brand identity the customer associates with it, we measure the amount of times users on X, formerly known as Twitter used both terms in a tweet as a percentage of total tweets. We study this at two distinct points in time, 4 months before and 4 months after the change.

1 GitHub

The public GitHub repo is available at Asartea/introduction-to-research-methods-final-project. It contains a zip file containing all source pdf's (`sources.zip`), a justification for the sources (`sources-justification.pdf`), the zip file of the overleaf project (`usage-of-twitter-vs-x.zip`, and the final report (`usage-of-twitter-vs-x.pdf`)

2 Introduction

In July 2023 Elon Musk renamed the social media platform Twitter to X. While parent companies of large brands do occasionally change their name, they usually do not touch the name of the actual products (When Facebook changed its name to Meta as an example, it only changed the name of its parent company; the name of the social media platform was and still is Facebook), out of fear of losing their build up brand identity. Musk not only changed the company's name but also jettisoned the Twitter branding and replaced it fully with the new X branding.

Branding is essential in fostering awareness of a service. Branding in the modern, digital era, is much more permanent than branding previously, when a customer's link with a brand would often be severed as soon as their transaction was completed (Edelman, 2010).

This rename has led to confusion about what the platform should now be called, and whether or not the term X is recognisable to the general audience. For this reason, the Associated Press Stylebook (APStylebook) recommends that writers introduce it on the first occurrence as "X, formerly known as Twitter", and only after that use the term X on its own (Conger, 2023). Similarly, the APStylebook also still allows the term tweet, and it is still widely used, despite X changing it to "post" in its official terminology.

In this study, we investigate to what degree this name change has succeeded in actually changing the term used by users on X, by looking at the percentage of tweets containing either of the terms at two points in time.

We hypothesise that due to the term Twitter being much more ingrained than any new term, while the number of tweets using X will have risen drastically the term Twitter will still be the dominant term used by users.

3 Related Work

As the topic is relatively recent, there is a lack of studies examining this specific area. Using Twitter for research into trends is a much studied area however, and the importance of brands has long been known.

Schroeder et al. (2019) discusses a Twitter search API based framework called FACT (Framework for Analysis and Capture of Twitter Graphs) for automatically retrieving tweets using certain keywords on a large scale, using an novel API-key sharing quota system to work around Twitter's rate limits on using the search API. FACT allows

a researcher to do complex work using network graphs, but also retrieve tweets based on specific keywords. It also lets the user search in the past, which allows for research which looks based on past events, allowing it to be more flexible in the types of research it supports.

Langguth et al. (2023) discusses the logistics and pitfalls of creating a keyword-based Twitter dataset using FACT, using it to search for all tweets containing keywords pertaining to COVID-19, and using that dataset to analyse the spread of COVID-19 conspiracy theories and misinformation. We use a similar, though much less complex method for this study.

James Andrew Kenyon and Bodet (2018) discusses the importance and mechanics of brand consistency, specifically in relation to the 2012 Olympics, but also theoretical background of the link people make with a brands identity.

Keller (1993) gives a further introduction to brand equity and how it develops, and presents a mental model of how a customer thinks about a brand. It also provides important definitions for brand related terms.

4 Data

To create the initial dataset we use the Twitter search API, targeting tweets and quote tweets. To work around Twitter's rate limiting and select data in the past we use FACT (Framework for Analysis and Capture of Twitter Graphs), first introduced by Schroeder et al. (2019). FACT allows us to use a distributed set of API quotas to search for specific keywords, and, importantly for this topic, look back in time.

Using FACT we select all tweets and quote tweets containing the keywords `Twitter` and/or `X`, sent between 0:00 UTC and 23:59 UTC on 23 March 2023, and between 0:00 - 23:59 UTC on 23 November 2023 (respectively 4 months before and after the change). We chose these dates because they provide a good snapshot of the number of tweets containing the keywords before and after the change, without getting too close to the date of the change itself, when we reasonably expect the number of mentions to temporarily have gone up massively as people discussed the news of the name change.

Pre-processing We randomly select 0.1% of all tweets in the dataset. We then filter out any tweet that contains a status other than *tweet* or *quote*, which removes all replies and retweets from the dataset. Including retweets would inflate the count for the terms, and we expect that users will mostly stick to the same term within the replies of a tweet. We also filter out any tweet that contains both the term Twitter and the term X.

Depending on whether or not this initial dataset shows that there is still a large number of tweets which use the term X at the first data point, further processing might be necessary to reduce the amount of false positives generated.[1]

We then tag all gathered tweets with three values: the date the tweet was send (which will be either March 23 or November 23), whether or not it contains the term "Twitter", and whether or not it contains the term "X" (Note that only one of those last two can be true for any given tweet). We then construct the desired percentage by dividing the amount of tweets which are sent on a date and contain the term by the amount of tweets that were sent in total on that day, and multiplying it by 100 to get a percentage. We repeat this process for all 4 date - term pairs to obtain our final data points.

Table 1 provides a summary of the data that will be used in this study.

Date	Contains Twitter?	Contains X?
Mar 23 or Nov 23	true/false	true/false

Table 1: Overview of the gathered dataset. For point in time we have gathered the percentage of tweets containing either term

5 Predicted Results

We expect that during the first data point, the usage of Twitter will be significantly higher than X and that the usage of X will be near zero. It is important to still check for the usage of X, however, in case there is another major source of tweets mentioning X. Such a source would necessitate adjusting the filtering done beforehand, before rerunning the experiment.

We expect that the usage of Twitter during the second data point will have fallen compared to the usage during the first data point, reflecting a major group of users switching terminology. We therefore also expect the usage of X to have risen in

[1] This is a limitation of carrying out this research only in the hypothetical. While we would expect the assertion that most tweets using X as a standalone word use it to refer to the social media platform, without actually examining the data this is impossible to prove.

comparison with the first data point. While there is likely some correlation between the two, as users switch terminology, investigating such a correlation is outside the scope of this study.

Table 2 summarises the obtained data as a 2d table showing the percentage of tweets containing either term split by date.

	March 23	November 23
Twitter	?%	?%
X	?%	?%

Table 2: Results split by term and point in time, with to be determined values marked as `?%`

Discussion This research shows that even when a brand owner radically switches the branding of a social media platform, including changing all its terminology, icons, etc, most users continue to use the familiar branding they have come to associate with a platform.

This is not entirely unexpected. When Musk renamed Twitter to X he thereby broke the link between it and the term Twitter, but that term had become ingrained in the public consciousness. As Rowley and Edmundson-Bird (2013) and Edelman (2010) point out, branding in the digital era has become more of a two way street, and it is not unexpected that a significant number of consumers has simply refused to change their mental model of what the branding is.

This presents an issue for any company who wishes to undertake such a re branding, and suggests that the approach by companies such as Meta (renaming only the company, not the platform itself) might be a better and easier understood way to change a companies branding, without generating the friction of renaming the platform itself.

6 Conclusion

This study aimed to investigate to what degree users on X have switched to using the term X instead of Twitter. We expect to find that while a significant minority has switched, a majority of users have not, and still use Twitter.

Further research could be done by focusing on analysing these same numbers on a week-by-week basis or even a day-by-day basis, which would provide a much clearer view of how this change develops over time. Two data points are only enough to see that there has been a change, but not see it develop.

Another research angle would be to explore to what degree it is possible to track users who moved from using Twitter to X through their journey, and see if they gradually start using one over the other or switch instantly.

References

Conger, K. (2023, Aug). So what do we call twitter now anyway?

Edelman, D. (2010, 12). Branding in the digital age: You're spending your money in all the wrong places. *Harvard Bus. Rev. 88*, 62–69.

James Andrew Kenyon, A. E. M. and G. Bodet (2018). Brand consistency and coherency at the london 2012 olympic games. *Journal of Strategic Marketing 26*(1), 6–18.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing 57*(1), 1–22.

Langguth, J., D. Schroeder, P. Filkuková, S. Brenner, J. Phillips, and K. Pogorelov (2023, 04). Coco: an annotated twitter dataset of covid-19 conspiracy theories. *Journal of Computational Social Science 6*, 443–484.

Rowley, J. and D. Edmundson-Bird (2013, 01). Brand presence in digital space. *Journal of Electronic Commerce in Organizations 11*, 63–78.

Schroeder, D. T., K. Pogorelov, and J. Langguth (2019). Fact: a framework for analysis and capture of twitter graphs. In *2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)*, pp. 134–141.