Celebrities: A Teaser for our new Holiday Paper Series

The Social Data Science Lab

December 26, 2024

In 2025, we will start our Summer Holiday Paper Series, releasing each year's paper in the Summer months. This Winter Holidays, we present this teaser of what is coming next summer. We are planning to study the popularity and public perception of celebrities with various methods. Here, we show some examples for Elon Musk, Mark Zuckerberg, and Britney Spears.

First, we did a survey simulation with GPT4 [1], where we asked the LLM to rate celebrities as if it was a person with certain demographic attributes in a particular year (2018, 2021, and 2024). We have no way to know if this is valid at all but had some expectations, so we tried with demographics in four combinations of age and gender (old-young, male-female) and contrasting queer with non-queer personas for the LLM. We asked for a few ratings on a 1 to 5 scale including ratings for "relatable", "relevant", "inspiring" and "likable".

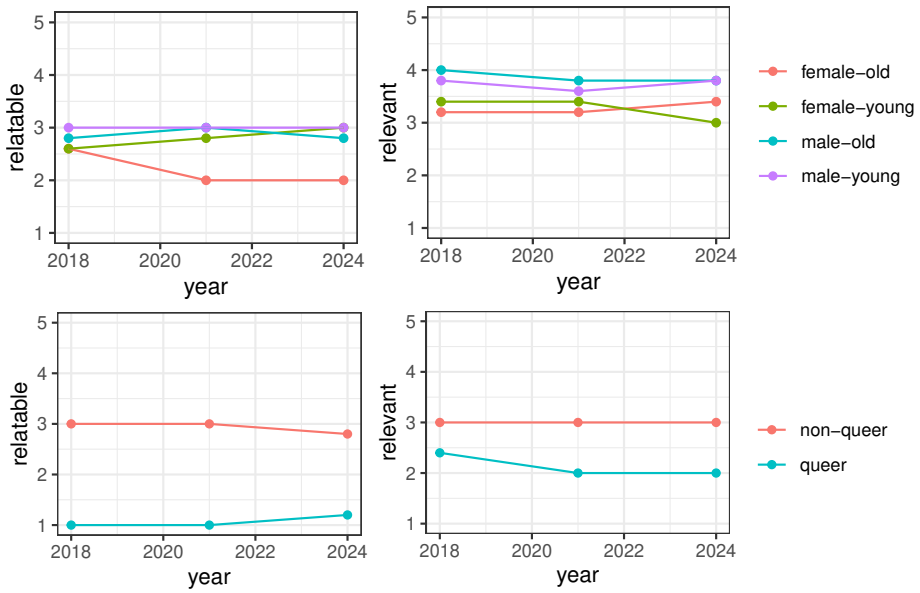

Figure 1: Simulated ratings for "relatable" and "relevant" for Elon Musk.

We find that LLMs with a female old persona rate Elon Musk (see Figure 1 with a slightly lower value on the relatable scale, but in general the ratings for Musk are rather stable over time and gender-age personas. Where there is a clear difference is with queer vs non-queer personas: Musk rates substantially lower with queer personas across ratings and this is stable over time.

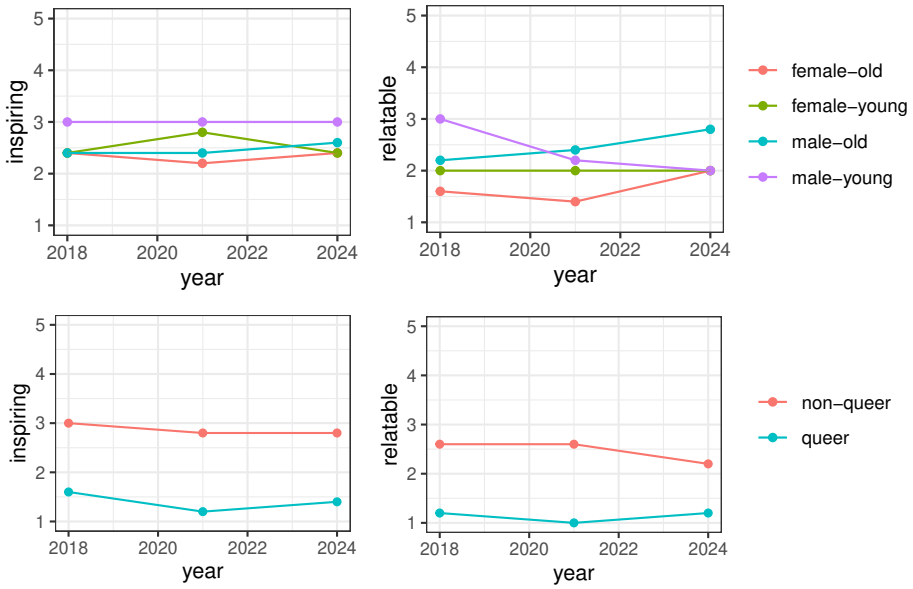

Figure 2: Simulated ratings for "inspiring" and "relatable" for Mark Zuckerberg.

Mark Zuckerberg is rated as pretty unrelatable, especially by the female old personas (Figure 2. Also Mark Zuckerberg rates much lower on scales (e.g. "inspiring" and "relatable") for queer personas as for non-queer ones.

We initially compared only Elon Musk and Mark Zuckerberg, but we wondered if the different levels of queer and non-queer personas was something that could be swapped for other celebrities. This is the case for Britney Spears, for whom queer personas rate much higher in "likable" and "inspiring" than non-queer personas. Across age and gender, female young personas are the ones that give Britney Spears the highest likability and male old personas the lowest relevance score.

To have something a bit more realistic than the above simulations of surveys, we turned to Reddit, our favorite social media data source nowadays. We analyzed all reddit comments up to Nov/2024 that contain terms related to Britney Spears, Elon Musk and Mark Zuckerberg ("elon", "musk", "zuck", "zuckerberg", "britney" and "godney"). We use an extensive filtering process to remove very short messages and spam, which removes both probable botted accounts and subreddits. This removed around half of all comments for each celebrity, resulting in 5.3M comments for Musk, 485K for Zucker-

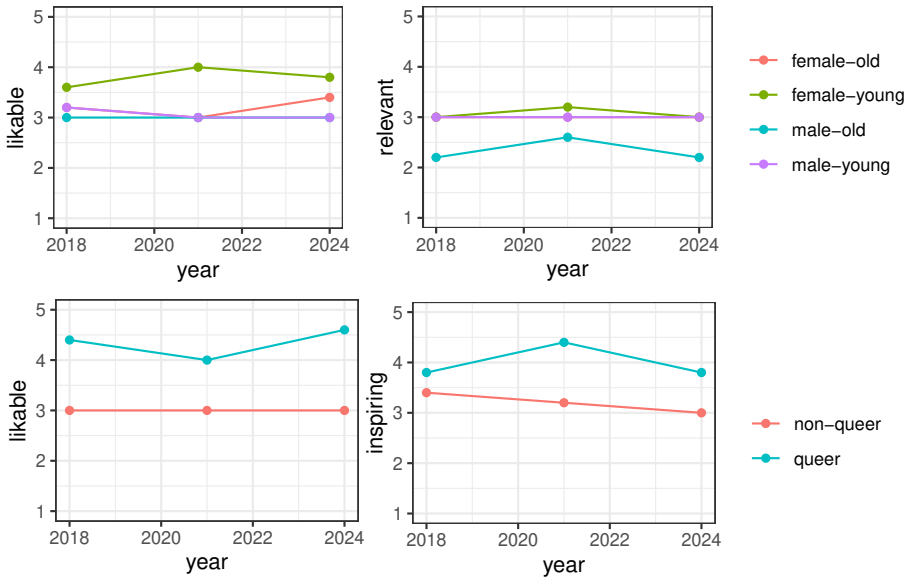

Figure 3: Simulated ratings for "relatable" and "inspiring" for Britney Spears.

berg, and 436K for Britney. We analyzed toxicity and emotions with optimized versions of *unitary/unbiased-toxic-roberta* (used in the detoxify library) and *SamLowe/roberta-base-go_emotions*, respectively. For toxicity we apply a threshold of 0.7, and for the emotions we use the optimized thresholds from the original model's page [2]. For the displayed timeseries, we use a 6-month rolling window. The full analysis code is available on Github [3].

We hand-picked three negative signals: share of toxic posts, share of posts expressing anger, and share of posts expressing annoyance, which can be seen with a 6-month rolling window on Figure 4. Comparing Musk and Zuckerberg, one can see that a decade ago, posts mentioning Zuckerberg were higher in toxicity, anger, and annoyance. But a clear trend for Musk made him surpass Zuckerberg in all three types of posts, where Zuckerberg has stayed mostly constant in comparison. For those of you disappointed with the fight that never happened between them, you can see here that Musk now loses a popularity contest, which would not be the case just a few years ago. Britney Spears is a good contrast, as all three types of posts have lower frequency that is stable over time.

On the time series of Elon Musk, you can see bumps in all three negative types of posts in the summer of 2018, coinciding with his fine by the SEC due to his market manipulation when tweeting about Tesla. What seemed to be peak unpopularity for Musk back then has clearly been beaten by recent years, as the current frequencies of negative posts are substantially higher than the peak of hate for Musk back in 2018.

To compare, we chose three types of positive posts expressing admiration,

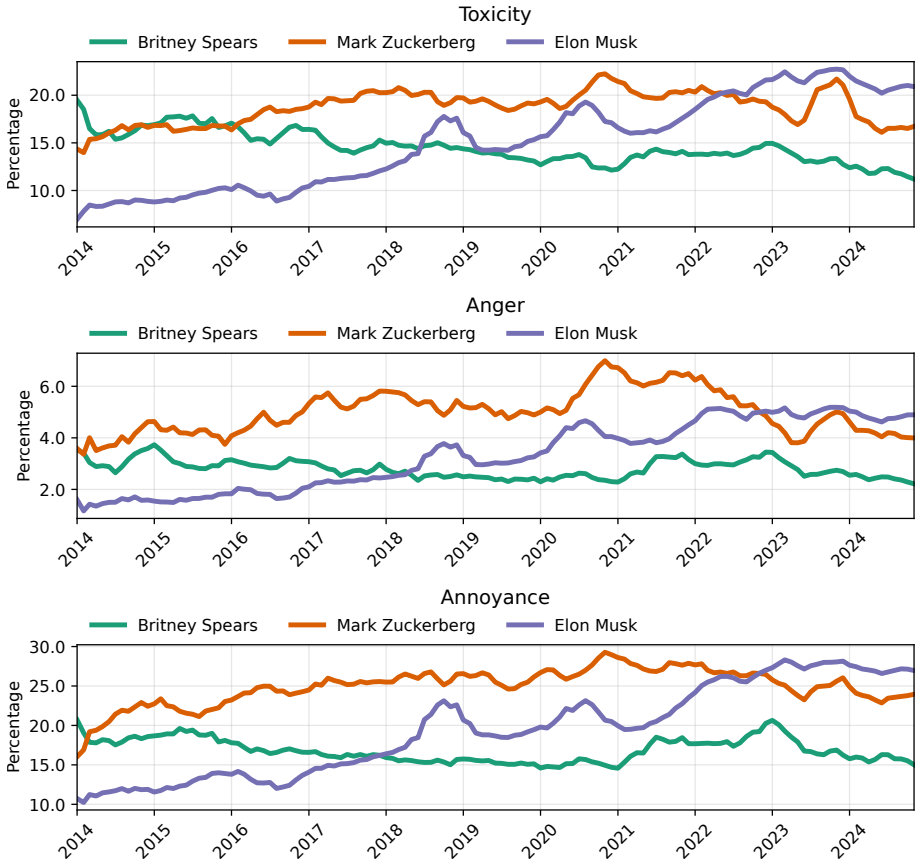

Figure 4: Time series of the percent of toxic posts and of posts expressing anger and annoyance for Elon Musk, Mark Zuckerberg, and Britney Spears.

approval, and excitement, shown on Figure 5. You can see that Britney Spears attracted much more frequent posts expressing admiration, having more than double the frequency of Mark Zuckerberg and Elon Musk in the recent years. The frequency for Zuckerberg has decreased between 2014 and 2017, while Musk had a steady decrease that has put him now in the same level as Zuckerberg.

Britney Spears never got much expression of approval, with a constant 8% of posts expressing it. Elon Musk and Mark Zuckerberg had much higher levels a decade ago, but in the recent years both managed to be slightly below Britney Spears. We wonder what could happen in a future US presidential election of Musk vs Spears.

In terms of excitement, Britney Spears has a steady level above the one of Mark Zuckerberg, but Elon Musk has moved from a high level similar to Britney Spears' to the level of Zuckerberg. One can say that nowadays, Reddit expresses more excitement when talking about Britney Spears than when taking about

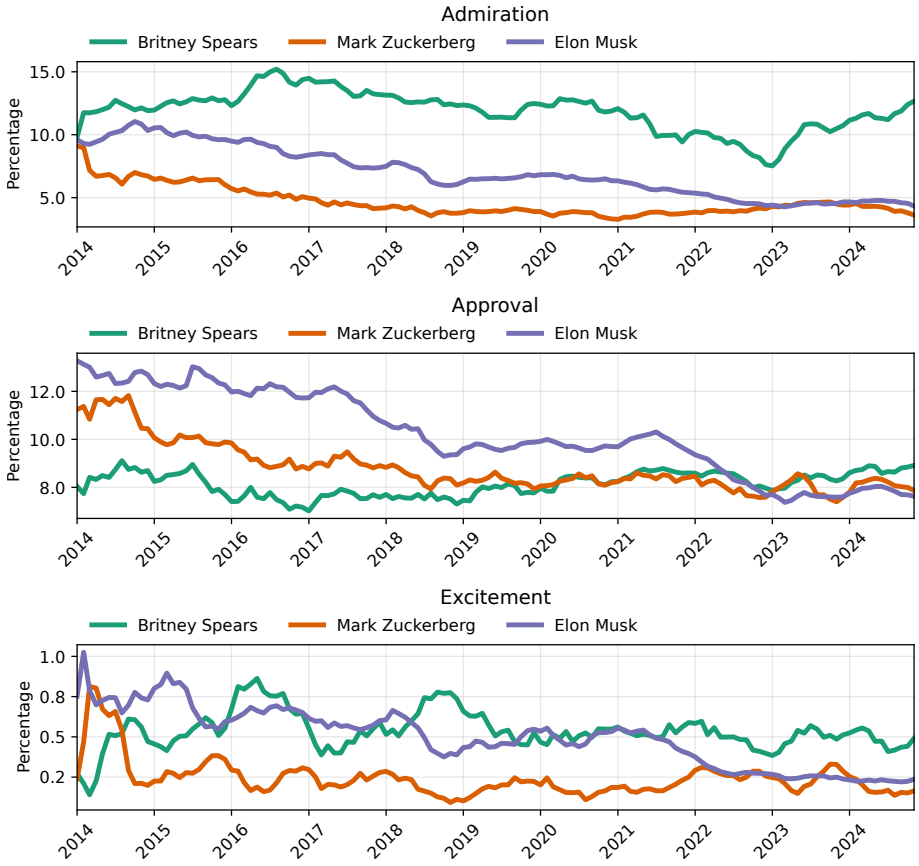

Figure 5: Time series of the percent of posts expressing admiration, approval, and excitement for Elon Musk, Mark Zuckerberg, and Britney Spears.

Musk or Zuckerberg. This mirrors what we found above for annoyance, as both Zuckerberg and Musk bring nearly double frequency of annoyance compared to Britney Spears.

This was a small teaser of what we are planning for the Summer. We will add more celebrities and study their life stories as reflected on Reddit and LLMs. We might even be able to explore the network of celebrities as they are co-mentioned on Reddit or gossip media or make popularity and unpopularity rankings alike. Stay tuned and happy Winter holidays!

References

[1] Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023.

[2] https://huggingface.co/SamLowe/roberta-base-go_emotions. Accessed: 2024-12-26.

[3] https://github.com/joaopn/christmas_paper_2024_reddit_celebrities. Accessed: 2024-12-26.