

I heart science: gendered language on Facebook

I wondered if there was a way to aggregate text conversations to build info-graphics. I was initially interested in my own text messages, wondering how many times i have used the word “dude”, for example. Or how many times my messages were about saying that I was on my way! The only downside was that I couldn’t find a way to export all of them as simple text. Once they were text I could easily parse them in Visual Basic. But I digress. I then thought about whether or not the same could be done with massive social network data.

I assumed that Google (the company itself) would do it in order to improve their search engines but I wondered if the information was available to little old me. It would be fascinating to convert that data in order to better grasp the “chatter”.

Turns out that [these researchers](#) did exactly that with data from Facebook. Fascinating. A lot of those gendered stereotypes are borne out in this study. The authors analyzed 700 million words, phrases, and topic instances collected from the Facebook messages of 75,000 volunteers, who also took standard personality tests, and found striking variations in language with personality, gender, and age.

So yes, there is a bias in the sample: these people are interested in participation. I wonder what everyone else is saying. I guess only the NSA can produce those infographics.

Source: <https://www.idonthaveacoolname.com/i-heart-science-gendered-language-on-facebook/>