**Web Scraping 4Chan /pol/**

After the recent [shootings](https://nypost.com/2019/03/14/new-zealand-mosque-shooter-livestreamed-killings-on-facebook/) in Christchurch, I became interested in conducting a quick discourse analysis on the website, 4chan and the group [/pol/](http://boards.4chan.org/pol/), which was said to be an online forum where the shooter was an active member.

When I first entered the forum, I felt a bit lost and overwhelmed mainly because of the amount of unstructured texts and images in it. I quickly realized that I would not be able to make sense of all this text just by reading them one by one.

I then decided to ask [R](https://cran.r-project.org/web/packages/rvest/rvest.pdf) for help. I expected it would help me grasp [keywords](https://cran.r-project.org/web/packages/tm/index.html), and perhaps a bit of the [tone](https://cran.r-project.org/web/packages/topicmodels/index.html) of this online environment. I then wrote down few [codes](https://github.com/ArgosMs/Web_Scrapping/blob/master/Web_Scrapping_4Chan_pol.R) to carry out this analysis.

The result emerging from the literal interpretation of the ten most frequently used words observed across eight pages, all posted in the spam of five minutes in the early afternoon of March 19th, 2019, suggest the following:

*The users of ‘pol’ are significantly concern about the future. They seem to make judgements based on similarities. They are extremely annoyed and are likely to make simplifications. Skin color, particularly the ‘white’ color, is a very popular topic among them. They are eager to promote political changes. They express criticisms through negation. They point the other as being different from themselves. They make judgements based on ‘rights’ or suspicion. They finally argue that they, or someone else, have to speak up.*

This discourse analysis provides a glimpse of the ‘tone’ used among /pol/ members. No grand claims can be made out of this. What this brief article does stress is that data science is here as a tool that help us understand the increasing complex world we live in. And those equipped with analytical and programing skills are likely to find the most adequate answers more efficiently.