WHAT DO WE TALK ABOUT WHEN WE PLAY PUNK!

A quantitative analysis of the lyrics of punk songs from the website www.plyrics.com

The data include 85,546 songs and 3,362 bands between 1977 to 2015. We extracted 12 topics from the corpus of songs. A "topic" is a pattern of words and phrases that occur together frequently across a set of songs. For each topic, we looked at the most characteristic words, the subgenres where the topic occurs more frequently, the sentiment expressed by the language, and the top 5 bands who talk about it in their lyrics. We also looked at the prevalence over time of every topic, to observe important differences which have been for the punk kids over the years. You can also see the time distribution of all topics together here on the right.

We gave a name to each topic based on the most characteristic words that you can see in the word clouds above. This is what each topic is defined as. To identify the subgenres, we used the 'genre' tags associated to the albums on plyrics. For many bands the genre is missing, so results are partial. The sentiment analysis is based on the 'NRC Word-Emotion Association Lexicon'.

The analysis was performed in R using the packages tm, wordcloud, phoemorpheme, tidynlp, and ggplot2.

Martin Mölder, once a member of a punk band himself, is now a doctoral candidate at the Politik Department at the Central European University. For more on him, see www.martinnolder.com

Federico Vegetti is a huge fan of punk rock since the age of 13, and works as a postdoc at the Politik Department at the Central European University. For more on him, see federicovegetti.github.io

Version: 2017.02.19