



Discovering Pandemic Topics on Twitter

Erik Tjong Kim Sang, Marijn Schraagen, Mehdi Dastani, Shihan Wang
Netherlands eScience Center / Utrecht University

2 June 2021



Approach

We search for topics in the pandemic discussion on Dutch Twitter. We use three methods for extracting topics from tweets:

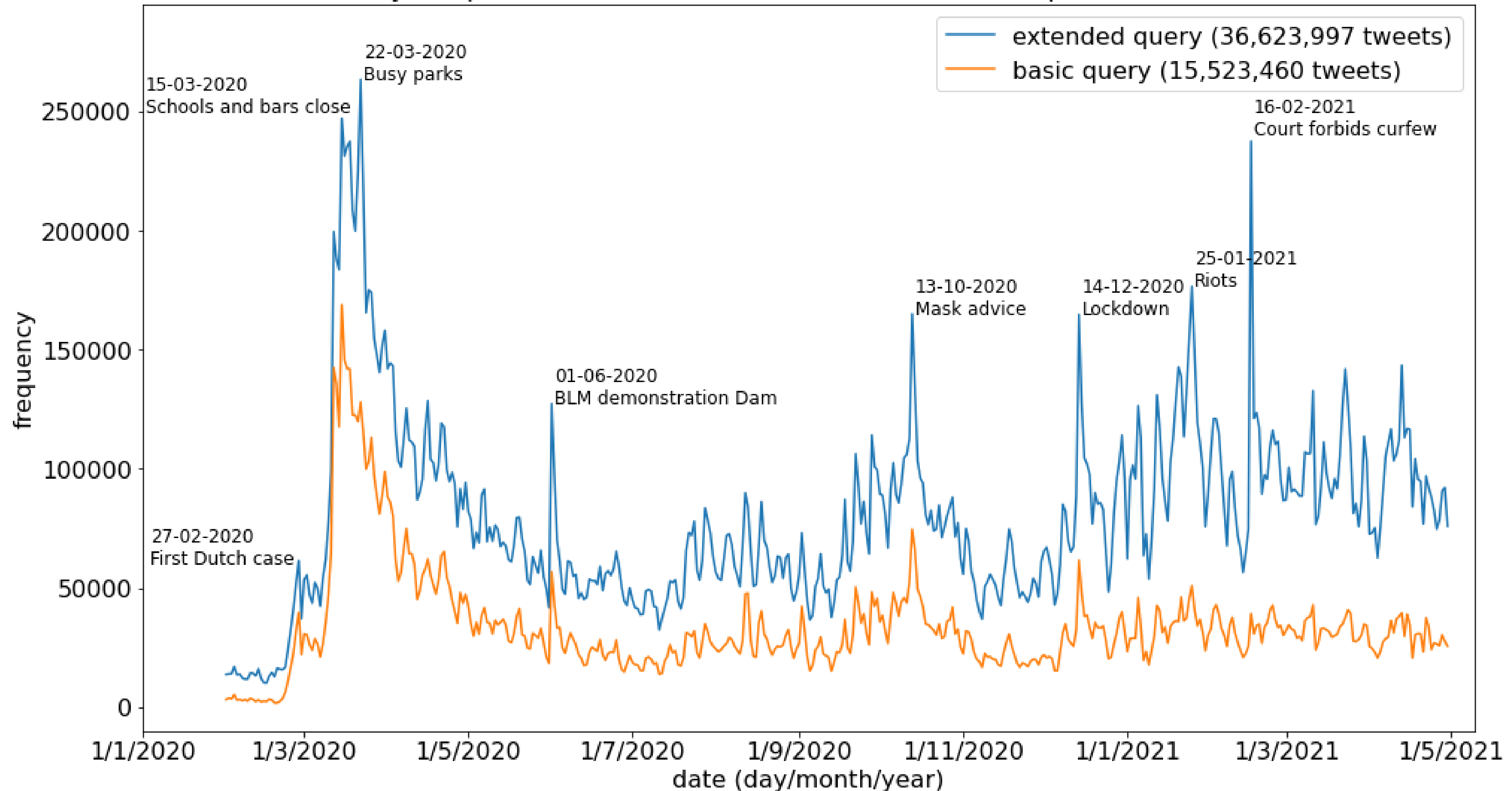
1. hashtag counts of tweets containing one of the strings covid or corona,
2. a statistical *t*-test comparison of the vocabulary used in these tweets and the vocabulary of the other tweets, and
3. Latent Dirichlet Allocation (LDA), a method for automatically extracting topics from texts

Results month January 2021

	hashtags	<i>t</i>-test	Latent Dirichlet Allocation
1.	measures	measures	riots
2.	pandemic	pandemic	pandemic
3.	curfew	vaccination	right-wing themes
4.	lockdown	curfew	vaccination
5.	FVD	lockdown	government
6.	vaccination	variants	schools
7.	protests	intensive care	curfew
8.	Rutte	testing	elections
9.	riots	FVD	well wishes
10.	press conference	hospitals	Urk

Application: extending search query for pandemic tweets

Daily frequencies of tweets written in Dutch with pandemic terms



Our paper with elaborate results: Erik Tjong Kim Sang, Marijn Schraagen, Mehdi Dastani and Shihan Wang, **Discovering Pandemic Topics on Twitter**. DHBenelux 2021. <https://ifarm.nl/erikt/papers/dhbenelux-20210509.pdf>



Our data (tweet ids): <http://145.100.59.103/downloads/>

Our other papers dealing with the pandemic on Twitter:

- Transfer Learning for Stance Analysis in COVID-19 Tweets, CLIN 2021
- Dutch General Public Reaction on Governmental COVID-19 Measures and Announcements in Twitter Data, Arxiv, 2020
- Public Sentiment on Governmental COVID-19 Measures in Dutch Social Media, EMNLP Workshop NLP for COVID-19 (Part 2), 2020

