Measuring public opinion in Russia through analysis of online petitions on change.org

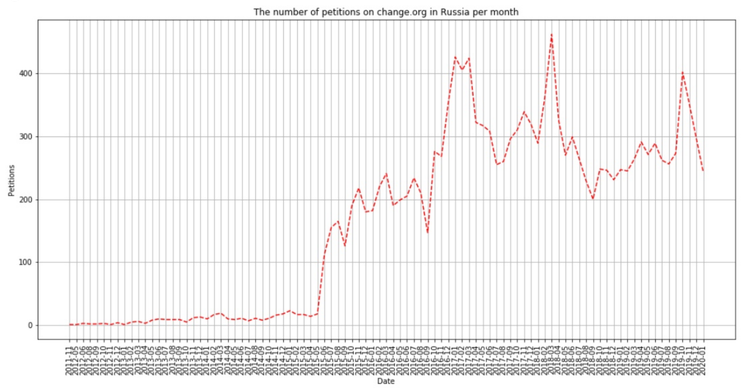
Polina REVINA
National Research University Higher School of Economics, Moscow, Russian Federation
pvrevina@edu.hse.ru — +7 (916) 038 35 26

Introduction

In this study, I try to interpret petitioners' priorities and explain the popularity of petitions, measured by a count of signatures on e-petitions. This analysis also indirectly points to latent topics of public opinion of Russian citizens that are not otherwise apparent and easily measurable. In authoritarian regimes, in the absence of effectively working formal institutions for interaction between citizens and politicians, platforms for creating online petitions become the main sources of engaging with authorities and expressing a public opinion[1]. One of the most popular projects for publishing petitions on the Russian internet is change.org. The ease of registration and use allows you to quickly create, distribute and promote e-petitions, thereby actively responding to events or adding new topics to the agenda[2].

Data

I collected 25000 petitions from 2011 to 2020, using a Selenium library for Python. The data structure includes the following variables: date of creation of the petition; addressees of the petition; a name of the petition; number of signatures under the petition; target number of signatures; petition's authors; a type of author (organization or owner); text of a petition; displayed petition tags.
The number of petitions on the website increased from only one in 2011 to roughly 5000 in 2019. Half of the petitions in the dataset are closed, and half are open, most petitions are created by organizations.



To reduce noise, petitions with text length and signature count of fewer than 50 were removed. I also applied preprocessing: remove stop-words, non-alphabetic and non-cyrillic characters, in the result, the number of petitions reduced to 15000.

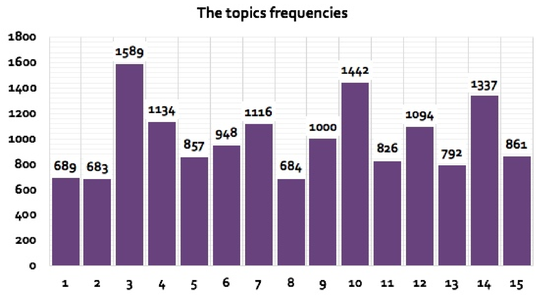
Method and model

Topic modelling with BigARTM

The topic model was built using BigARTM library for Python that is based on the principle of additive regularization. Compared to LDA and PLDA algorithms, BigARTM performs more flexible strategies to derive topics from large-scale text corpora. Confidence of the model in one topic by more than 0.5 corresponds to 67% of petitions. According to the results of the model, I identified the following topics:

Topic	Tokens
Healthcare system	doctor, treatment, medical, hospital, drug, patient, health, death
Housing problems	house, apartment, builder, construction, contract, tariff, facility
Medical care for children	child, help, life, family, son, live, why, nobody, become, time
Territory development	district, construction, resident, city, territory, urban, development
Social system institutions	Kindergarten, disabled person, social, help, institution, right, child
Civil rights	Russian, law, citizen, country, authority, people, freedom, society
Credits and loans	money, ruble, bank, tax, account, pay, cost, credit, income, debt
Historical heritage	monument, historical, unique, cultural, natural, heritage, time
Science and education	school, education, director, work, student, leader, ask, science
Transport infrastructure	city, road, street, transport, car, traffic, driver, bus, repair, hour
Ecology	water, plant, region, ecological, waste, garbage, territory, river
Patriotism	ours, Russia, own, country, peace, war, great, language, become, time
Regional identity	region, ours, Russia, own, problem, governor, budget, center, official
Animal rights	animal, dog, shelter, homeless, cruel, animal, zoo, kill, condition
Criminal prosecution	court, case, criminal, judicial, decision, crime, investigation, action

The most common topics among petitions are medical care for children, transport infrastructure, animal rights and territory development, and least of all petitions is dedicated to healthcare, housing problems and preservation of natural and historical heritage. Moreover, the popularity of most of the topics did not change over time, only the quantity of petitions about criminal prosecution and civil rights raised significantly in 2017. So, I can conclude that these are crucial points that form a public opinion of Russian citizens.



Prediction of petitions' popularity

Next, I estimated a linear regression model to predict petitions' popularity. I logged number of petition signatures due to its skewed distribution. Besides topics, which were described above, I added the following variables to the model:

- **A type of petition's owner**: user or organization. It is assumed that the organization has more resources for distributing the petition, therefore, a positive relationship between the type of the owner and target variable will be observed.
- **Number of names** and **geographical objects detected in a petition**, which were obtained using rule-based named entity recognition library for the Russian language – Natasha. It is assumed that large number of named entities in the text indicates more specificity and details, which can be positively associated with the popularity of the petition.

Results

According to the results of the regression model, only the independent variable for Topic 2 is significant. So, if a petition is concerning housing problems, it will increase the logged number of petition signatures by 0.282. Number of geographical features is not significant, however, an increase in a number of names mentioned in a petition leads to an increase in the logged number of signatures by 0.014. If the petition is written by an organization and not by an ordinary user, then the value of the dependent variable is increased by 0.143.

Thereby, I highlighted the main topics found in Russian-language petitions on change.org, examined their popularity and dynamics over time, thus measuring the public opinion of Russians. I also examined how these topics and other features, can predict the popularity of a petition. As a result, it is assumed that not so much the topic as the way they are distributed and detailed are connected to petitions' popularity.

References

[1] Andrei V Chugunov, Yury Kabanov, and Ksenia Zenchenkova. Russian e-petitions portal: exploring regional variance in use. In *International conference on electronic participation*, pages 109–122. Springer, 2016.

[2] Ahmed Said Elnoshokaty, Shuyuan Deng, and Dong-Heon Kwak. Success factors of online petitions: Evidence from change. org. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 1979–1985. IEEE, 2016.