

# Voices and Responses

---

A Large-Scale Text Classification of  
Grievance Forums in Autocracies

Aidar Zinnatullin

Postdoc at IKM (RPTU)



# Research Focus: Constituency Service / Responsiveness in the Authoritarian Setting

---

- > Two grievance channels in autocracies
  1. Opposition-provided infrastructure
  2. 'Good Governance' practices of autocrats

# Research Focus: Constituency Service / Responsiveness in the Authoritarian Setting

---

- > Two grievance channels in autocracies
  1. Opposition-provided infrastructure
  2. 'Good Governance' practices of autocrats
- > Under what conditions do citizens turn to the state-promoted platforms to seek redress?

# Research Focus: Constituency Service / Responsiveness in the Authoritarian Setting

---

- > Two grievance channels in autocracies
  - 1. Opposition-provided infrastructure
  - 2. 'Good Governance' practices of autocrats
- > Under what conditions do citizens turn to the state-promoted platforms to seek redress?
- > Do citizens *substitute* the low-risk digital channel provided by the authorities for high-risk protest when repression rises?

# Research Focus: Constituency Service / Responsiveness in the Authoritarian Setting

---

- > Two grievance channels in autocracies
  1. Opposition-provided infrastructure
  2. 'Good Governance' practices of autocrats
- > Under what conditions do citizens turn to the state-promoted platforms to seek redress?
- > Do citizens *substitute* the low-risk digital channel provided by the authorities for high-risk protest when repression rises?
- > Are state's responses performative or substantive?

# Where to Find Citizens' Complaints?

Putin called regional authorities to organize effective communication mechanisms with citizens with the use of new technologies: <http://kremlin.ru/events/president/news/62701>; English version: <http://en.kremlin.ru/events/president/news/62701/photos>



## Visit to Regional Governance Centre

Vladimir Putin visited the Regional Governance Centre in the Moscow Region Government building. The President familiarized himself with the digital solutions for urban management and communication with local residents.

January 30, 2020 18:30 Krasnogorsk 6 photos



# Where to Find Citizens' Complaints?

---

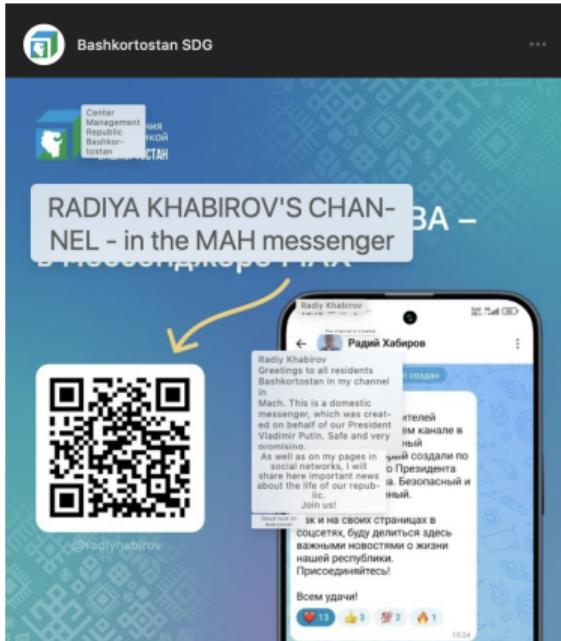
1. All high ranking officials and state organizations at a regional/local levels have to create and operate their accounts on social media platforms

# Where to Find Citizens' Complaints?

---

1. All high ranking officials and state organizations at a regional/local levels have to create and operate their accounts on social media platforms
2. The idea is 'to go to the people' to strengthen the trust of the public and convince them that authorities are competent

# Illustration 1



Radiy Khabirov created a channel in the MAX messenger. This is a domestic messenger, which was launched on behalf of the President of Russia.

The channel will publish important news about the life of Bashkortostan.

You can subscribe by following the link: [max.ru/radiykhabirov](http://max.ru/radiykhabirov)

Like 21 Share 5 23 hours ago

14

# Illustration 2

First the old ones -

**Aider Tukhbatulin**  
Hello  
Why is there no light in the fish, all the streets are sitting at night. I wrote that the machine burned out, it's been a week, they don't do anything. And now it's getting dark early.  
yesterday at 16:34 Reply Share

**Tymazinsky district of the Republic of Bashkortostan**  
Aider Tukhbatulin, hello. After the heavy rains that were last week, there was an emergency with the supply of power supply. They can't find the reason yet. In the near future, specialists from the resource supply organization will come for the examination. The question is under the control of the administration of the JV Kakyrbayhevsky Village Council. Elena Salimova, an employee of the administration of Tymazinsky district.  
today at 7:19 Reply Share

**Anna Mikhalova**  
Hello! Bicycle and pedestrian path from the intersection of St. Kuyukina-Rabkorov to the intersection of Rabkorov - Mubayakov and further overgrown, it is impossible to pass or drive. Please take action.  
Photos are attached.  
  
yesterday at 17:01 Reply Share

**Administration of Kirovsky district**  
Hello, Anna. Your message about the crowning of trees on the sidewalk from Rabkorov Street to Stepan Kuyukin Street will be worked out with the district services. Nikita Matveev, an employee of the administration of the Kirovsky district of the city Ufa.  
today at 6:21 Reply Share

**Anna Vladimirova**  
Please help me in this situation. We are tired of such conditions.  
  
yesterday at 18:13 Reply Share

**SOG Bashkortostan - Author**  
Hello, Anna. Please specify the settlement and address of the house, Reziya Musibirova, employee of the Management Center of the Republic of Bashkortostan  
yesterday at 18:44 Reply Share

**Oktjabrsky district of Ufa**  
Hello, Azamat Zagirov. Colleagues from the specialized services were given information about the need to take out curbs on the territory of the school № 114. The removal was scheduled before the end of the month. Let's add that outsiders are prohibited from staying on the territory of educational institutions. If you have any questions, you can contact the education department of the municipal administration by phone 234-41-55. Gusal Adilov, head of the education department of the administration of the Oktjabrsky district of Ufa.  
today at 7:12 Reply Share

**Julianna Verne**  
Due to the absence of a school in Kuznetsovsky Zaton, we ask you to speed up the delivery of Pugachev<sup>IT</sup> to date, there is no work on the section from Sochi to Rybko, there have been no workers, no equipment since the beginning of August 2025. Children should be taken to school from September 1. Less than 2 weeks left!  
  
today at 5:53 Reply Share

**#Kirovsky district administration, @ratimir.mavilev, @kiradimula, @radiybaibarov, @gavbash, @alexander\_bastrikin, #Putin**  
  
today at 5:53 Reply Share

**Miliya Timer**  
Julianna, yes... it's the middle of August 2025  
  
Video from Miliya Timer  
3 views today at 6:23 Reply Share

**Larisa Yusupova**  
It is because of the protracted construction of the road to Pugachev that residents cannot get to educational institutions and places of work in a timely manner. That is why it is not possible to let in additional buses, so that citizens of the republic living in Kuznetsovsky Zaton do not stand in debt at stops, not being able to leave quickly and not stand in the rain or in the cold.  
today at 6:03 Reply Share

**Ritvaly Yabov**  
Hello, the order for new routes for Kuznetsovsky Zato-zato, which was given to the Ministry of Transport on 04.08.2025, at a meeting in the

**Vladimir Korolik**  
And what's the use of the fact that the channel of the head of the REPUBLIC was created, we ordinary residents of the city of Nevezino have not been heard. Here in the city of Yanau Nekrasova street number 17, since April, can't be restored laundry drying platforms. People are forced to walk on such stones and dirt. Officials respond that we keep the situation under control Show More  
  
today at 7:34 Reply Share

**Administration of Yanau district of the Republic of Belarus**  
Vladimir Korolik, hello. All your questions have been repeatedly explained:  
[https://vk.com/wa415272232\\_14534419](https://vk.com/wa415272232_14534419)  
reply=1454169..., [https://vk.com/wall-144352066\\_298263?reply=298298...](https://vk.com/wall-144352066_298263?reply=298298...),  
[https://vk.com/wall-20146046\\_1678147](https://vk.com/wall-20146046_1678147)  
reply=167864, etc. The situation has not changed to this day. Restoration of the coating of the laundry drying area in the yard of the house at 17 Nekrasova Street will be carried out until 01.09.2025, as planned. Regarding the placement of laundry drying areas in the center of the park, it was emphasized that it is not provided for by the improvement project. As for safety at the pedestrian crossing on 21 Nekrasov Street. A video camera will be installed at this site in the near future. We will report the results later. Olga Ismagilova, chief specialist of the information and analytical department of the administration of the Yanau district.  
today at 8:09 Reply Share

**RPTU**

# Data Collection: Puppeteer as Alternative to Selenium

- > Automated collection of full-page HTML snapshots from a prepared list of URLs using an authenticated browser session: <https://pptr.dev>
- > Simulation: launch browser → set viewport → open vk.com → type VK username submit → confirm login → iterate CSV links → for each: open URL, wait for network idle, capture full DOM HTML, save to disk, pause 3s → close browser
- > simple parsing with 'rvest', as for static webpages
- > another alternative is Playwright <https://playwright.dev>

Version: 24.19.0

## Puppeteer

Puppeteer is a JavaScript library which provides a high-level API to control Chrome or Firefox over the DevTools Protocol or WebDriver BIDI. Puppeteer runs in the headless (no visible UI) by default

[Get started](#) | [API](#) | [FAQ](#) | [Contributing](#) | [Troubleshooting](#)

### Installation

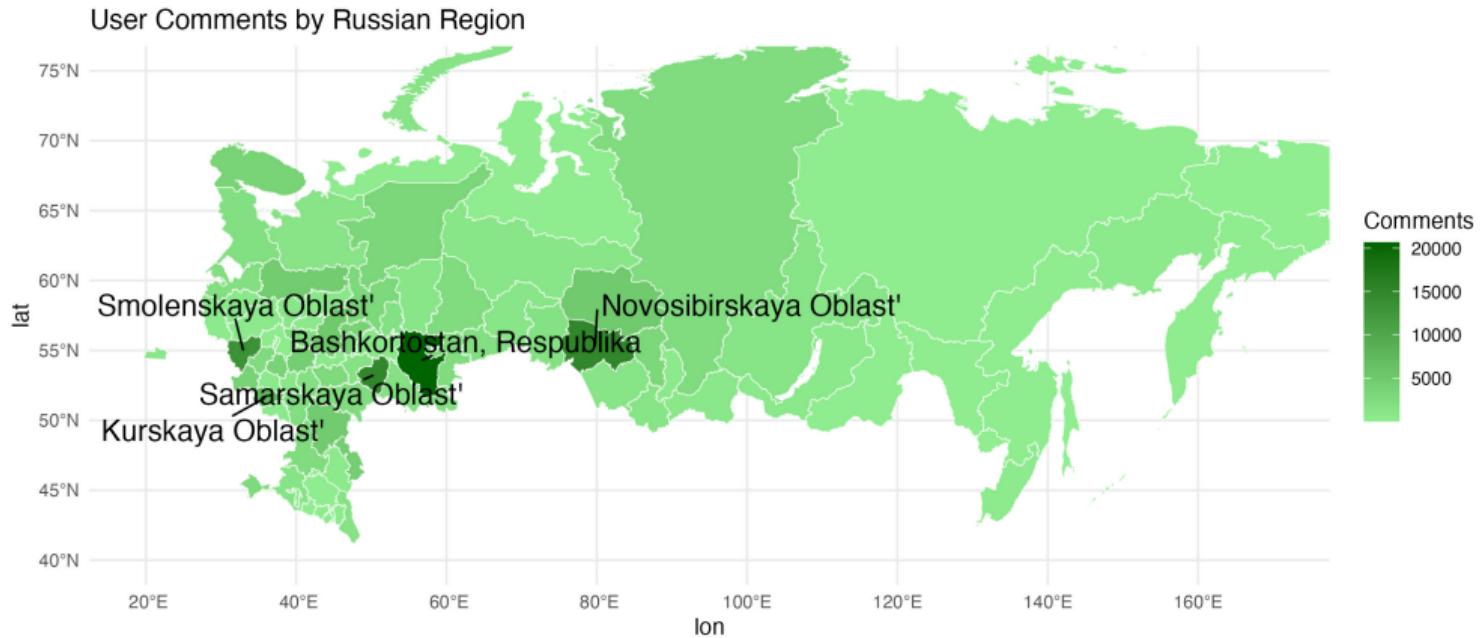
[npm](#)   [Yarn](#)   [pnpm](#)

```
npm i puppeteer # Downloads compatible Chrome during installation.  
npm i puppeteer-core # Alternatively, install as a library, without downloading Chrome.
```



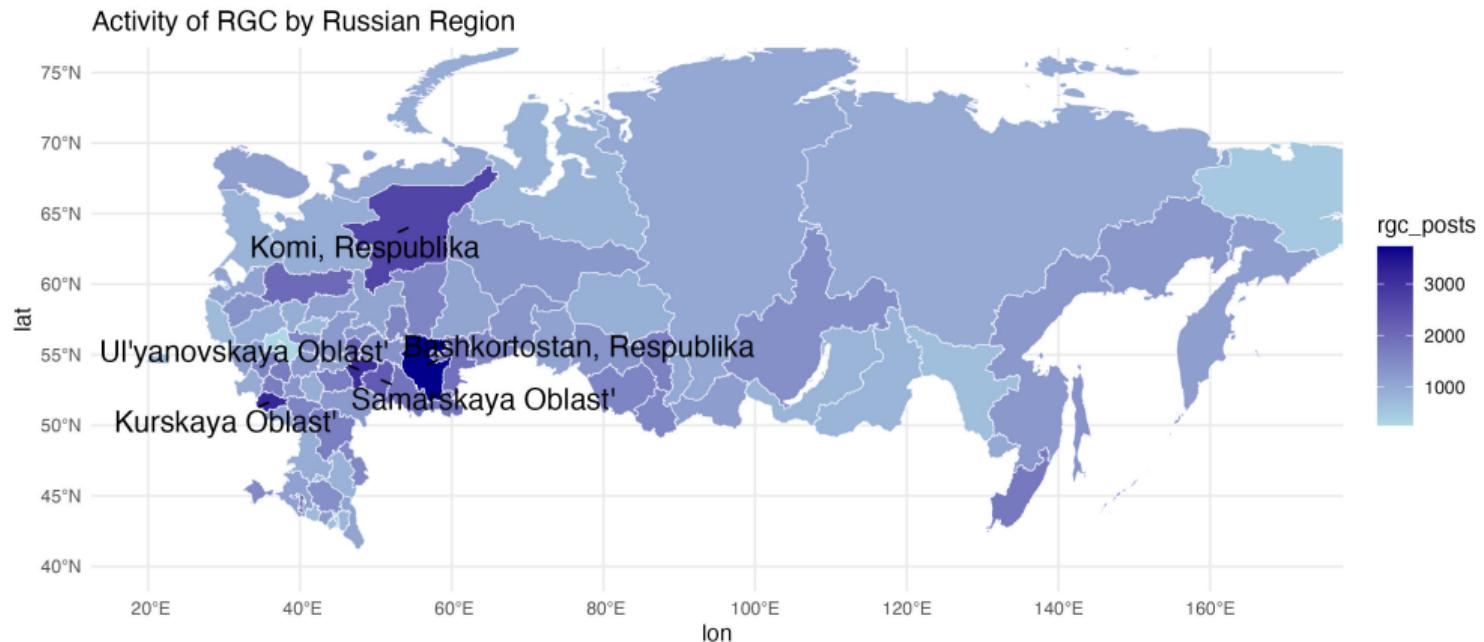
# Regional Distribution of User Comments

---

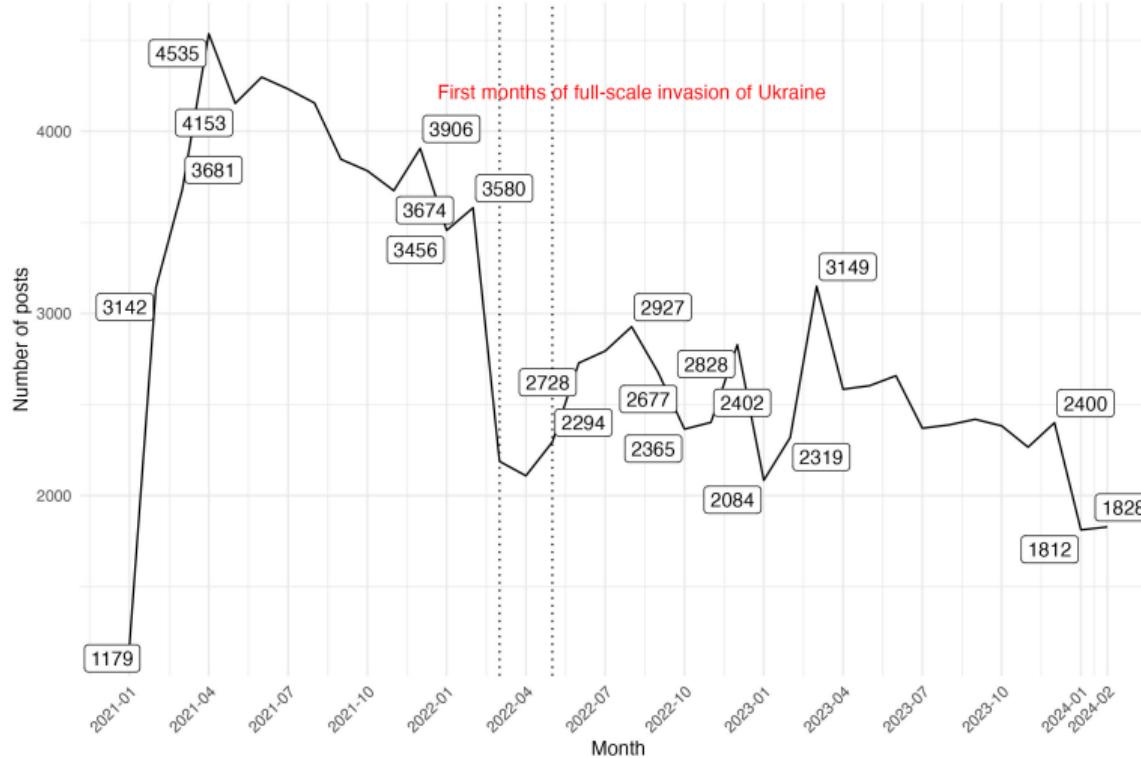


# Regional Distribution of RGC Activity

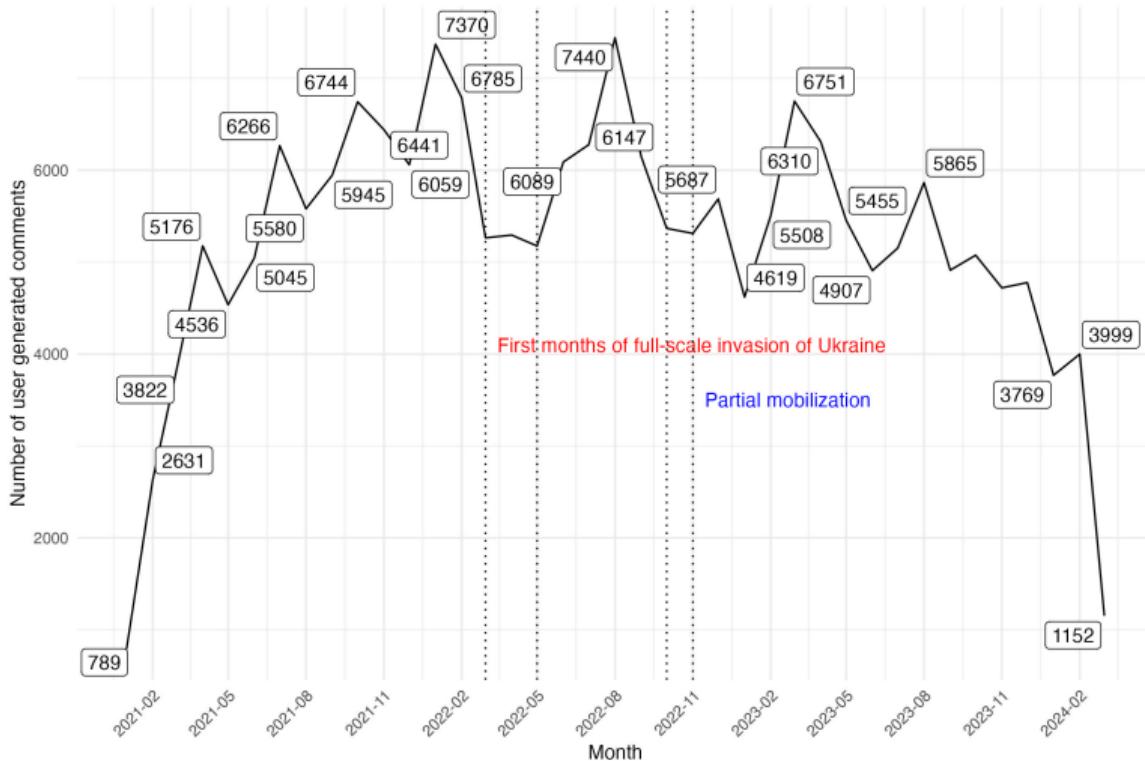
---



# Distribution over Time, RGC Posts



# Distribution over Time, User Comments



# Identifying Citizens' Digital Grievances on VK

---

1. **Focus:** VK communities run by Russia's *Regional Governance Centers* (RGCs) that report on citizen-reported issues
2. **Snowballing approach** for official pages
  - > seed keywords → candidate pages
  - > page categorized *official* if it met any 2 of 4 cues: (1) affiliation in header, (2) verified badge, (3) territorial scope, (4) authoritative replies
  - > iterative process → stopped after 4 waves produced no new officials
3. **Corpus**
  - > 215 596 comments by official accounts
  - > 233 048 comments by ordinary users
  - > comments posted by ordinary users < 5 words were removed (emoji, "thanks!", etc.). Therefore, the corpus of user comments was reduced to 204,530 comments
4. **Validation**
  - > Krippendorff's  $\alpha = 1.0$  on 300 double-coded accounts
  - > Automated check: macro F1 = 0.93
  - > Wrong labels are mostly officials using personal profiles

# Accounts Classification

---

- > Official accounts identified through snowball sampling (215 596 official comments, 290 545 user comments).
- > Krippendorff's  $\alpha = 1.0$  (perfect reliability).

Category	Precision	Recall	F1 Score
User	0.91	0.94	0.93
Official	0.95	0.91	0.93
Overall Accuracy			0.93
Macro F1 Score			0.93

Tabelle: Performance metrics for accounts classification

# Structural Topic Modeling (STM)

---

## 1. Pre-processing

- > tokenization; lowercasing; remove punctuation, stopwords ( $n = 563$ ), special characters, emojis, numbers
- > lemmatization with Yandex MyStem (better than stemming for Russian; May, Cotterell, & Durme, 2016)

## 2. Choosing $k$ (Roberts et al., 2019; searchK)

- > coarse search  $k = 10\text{--}100$  (step 10) → refine  $10\text{--}40$  (step 5) → final  $11\text{--}20$
- > selected  $k = 21$ : best balance of *semantic coherence & exclusivity*

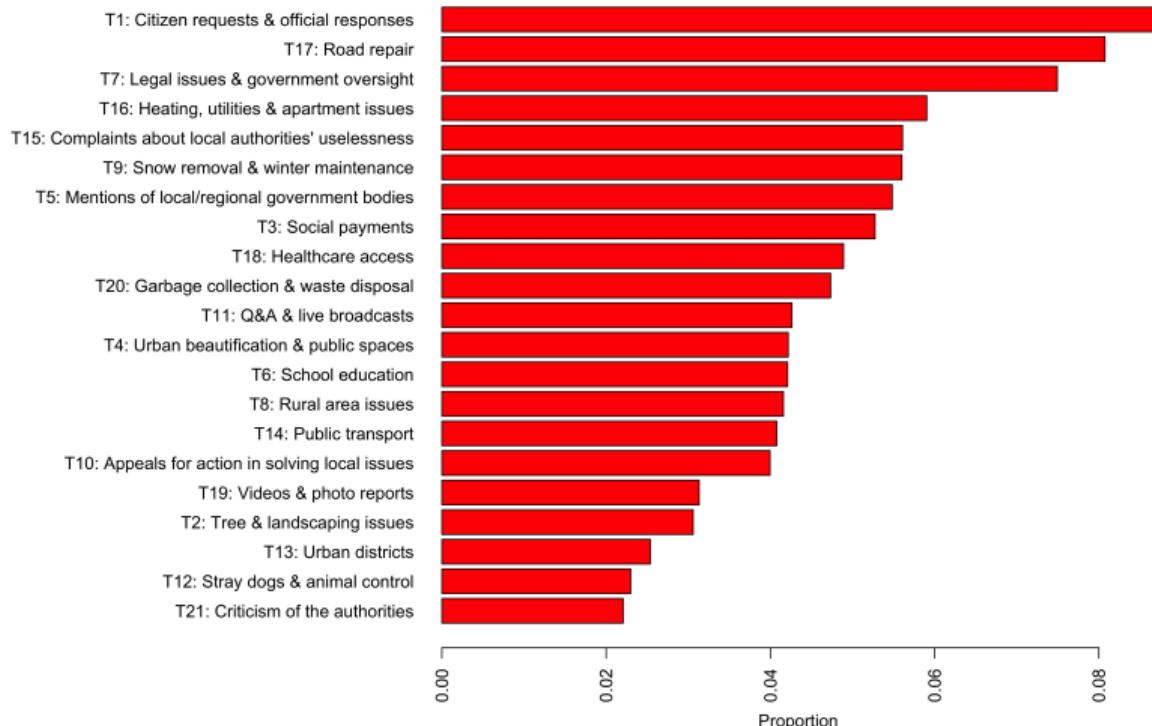
## 3. Human validation (Chang et al., 2009; oolong: Chan & Sältzer, 2020)

- > word intrusion: mean precision = 0.83,  $p < 0.001$  – coherent word groupings
- > topic intrusion: mean TLO =  $-1.94$ ,  $p = 0.976$  – overlap because of the nature of the comments, i.e., complaints
- > reliability: Krippendorff's  $\alpha = 0.498$  – moderate, adequate for exploratory aims

# What are People Posting?

---

Topic proportions, STM for k = 21

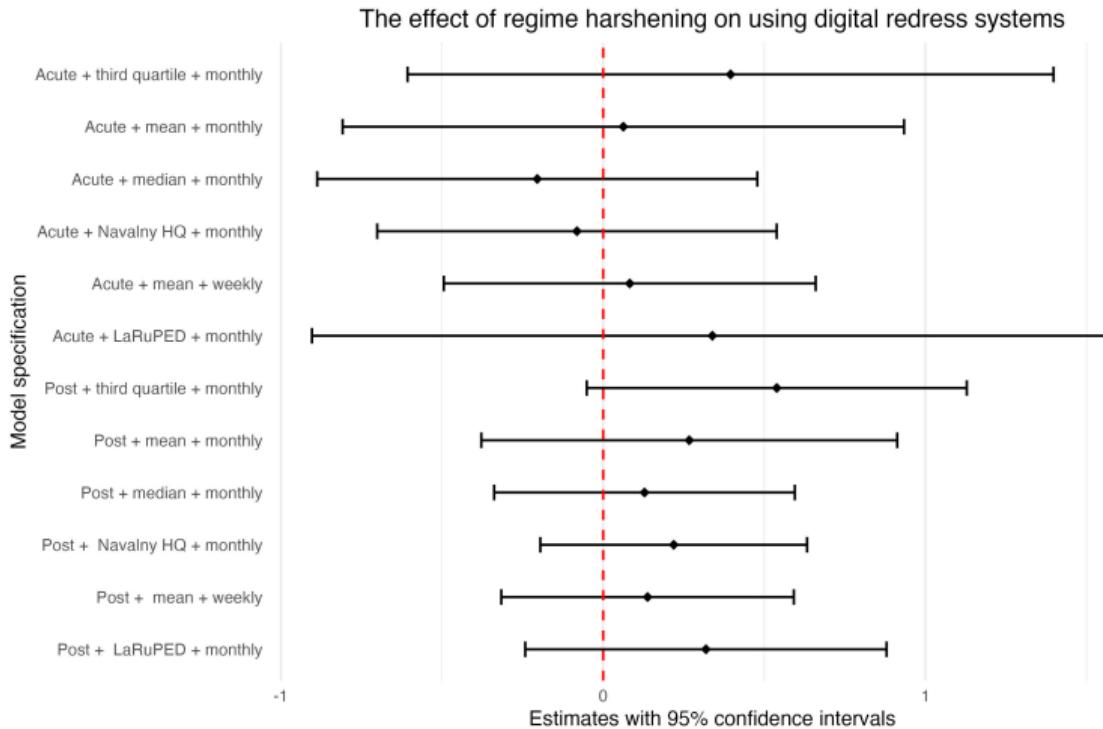


# H1 ( $\uparrow$ Protest $\implies$ $\uparrow$ Digital Complaints)

Variable	Model 1 (main)	Model 2 (no outliers)	Model 3 (two-way FE)	Model 4 (parsimonious)
Intercept	0.998 (2.475)	1.398 (2.477)	-1.835 (2.477)	4.726*** (0.229)
Protest Activity	0.431*** (0.108)	0.409*** (0.118)	0.490*** (0.098)	0.567*** (0.105)
RGC Activity	0.865*** (0.161)	0.770*** (0.165)	1.135*** (0.151)	-
UrbanShare	0.010 (0.007)	0.012+ (0.007)	0.011 (0.008)	-
RetiredShare	0.028 (0.027)	0.032 (0.027)	0.037 (0.028)	-
Internet	-0.032+ (0.017)	-0.033+ (0.018)	-0.035* (0.017)	-
Unemployment	-0.024 (0.043)	-0.021 (0.042)	-0.019 (0.037)	-
Region FE	No	No	Yes	No
Year FE	No	No	Yes	No
Observations	252	246	252	252
R <sup>2</sup>	.311	.290	.384	.123

Notes: Standard errors in parentheses. + $p < 0.10$ , \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

## H2 (Post-Invasion Replacement of Protest by Digital Redress Systems)



PA

## Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI

Moritz Laurer<sup>✉</sup>, Wouter van Atteveldt<sup>✉</sup>, Andreu Casas and Kasper Welbers

*Department of Communication Science, Vrije Universiteit Amsterdam, Amsterdam, Netherlands. Email: m.laurer@vu.nl, wouter.van.atteveldt@vu.nl, a.casassalleras@vu.nl, k.welbers@vu.nl*

---

### Abstract

Supervised machine learning is an increasingly popular tool for analyzing large political text corpora. The main disadvantage of supervised machine learning is the need for thousands of manually annotated training data points. This issue is particularly important in the social sciences where most new research questions require new training data for a new task tailored to the specific research question. This paper analyses how deep transfer learning can help address this challenge by accumulating "prior knowledge" in language models. Models like BERT can learn statistical language patterns through pre-training ("language knowledge"), and reliance on task-specific data can be reduced by training on universal tasks like natural language inference (NLI; "task knowledge"). We demonstrate the benefits of transfer learning on a wide range of eight tasks. Across these eight tasks, our BERT-NLI model fine-tuned on 100 to 2,500 texts performs on average 10.7 to 18.3 percentage points better than classical models without transfer learning. Our study indicates that BERT-NLI fine-tuned on 500 texts achieves similar performance as classical models trained on around 5,000 texts. Moreover, we show that transfer learning works particularly well on imbalanced data. We conclude by discussing limitations of transfer learning and by outlining new opportunities for political science research.

*Keywords:* machine learning, computational methods, text as data, transfer learning

---

# NLI Example for Constituency Service / Responsiveness

---

Given hypothesis  $H$ : “The state organization resolved my problem.”

Hypothesis	Context	NLI Class
The state organization resolved my problem.	“Your case has been resolved. The updated address is now in our records. No further action is required.”	True
The state organization resolved my problem.	“Your request is under review. Estimated completion: 10 business days.”	False
The state organization resolved my problem.	“Public notice: our office hours have changed to 08:00–16:00.”	False

Run the same test with  $H_2$ : “The state organization is still working on my problem.” and  $H_3$ : “The state organization needs more information or is giving general information.”

Pick the label whose hypothesis is most *entailed*:  $H_1 \rightarrow 1$  (Solved),  $H_2 \rightarrow 2$  (WIP),  $H_3 \rightarrow 3$  (Need info / General).

# Training Arguments for NLI

---

- > Training set contained 300 comments, balanced dataset.
- > A human annotator validated the classification (two-week interval, intracoder check Krippendorff's  $\alpha = 0.925$ ).

```
model_name = "cointegrated/rubert-base-cased-nli-twoway"
training_args = TrainingArguments(
    output_dir="output_nli_russian",
    evaluation_strategy="no",
    per_device_train_batch_size=8,
    per_device_eval_batch_size=8,
    num_train_epochs=3,
    learning_rate=2e-5,
    logging_steps=50,
    save_steps=500)
```

# Performance of NLI (First Round)

---

Category	Precision	Recall	F1 Score
Solved Situation	0.83	0.85	0.84
Work in Progress	0.81	0.85	0.83
Other	0.84	0.78	0.81
Overall Accuracy			0.83
Macro F1 Score			0.82

Tabelle: Performance metrics for the NLI classification, first round

# Performance of NLI (Second Round)

---

Category	Precision	Recall	F1 Score
Solved Situation	0.79	0.85	0.82
Work in Progress	0.74	0.80	0.77
Other	0.79	0.68	0.73
Overall Accuracy			0.78
Macro F1 Score			0.77

Tabelle: Performance metrics for the NLI classification, second round

# Results, Classification of State-related Comments

---

<b>Comment type</b>	<b>N of comments</b>	<b>Share (%)</b>
Solved	18,732	9%
Work in Progress	76,511	35%
Other	120,353	56%

Tabelle: State-related comments overview

# Next Steps

---

- > Analyse content of citizen–official threads to map *issue types* and regional/thematic variation

## Next Steps

---

- > Analyse content of citizen–official threads to map *issue types* and regional/thematic variation
- > Examine whether higher RGC posting = higher *resource allocation* for problem solving

## Next Steps

---

- > Analyse content of citizen–official threads to map *issue types* and regional/thematic variation
- > Examine whether higher RGC posting = higher *resource allocation* for problem solving
- > Cross-country / cross-regime comparisons

> Thank you!

> Thank you!  
> aidar.zinnatullin@rptu.de