

Framing and Topic Modelling in Belarusian State-Owned Media

Tsimafei Prakapenka

Saarland University

Dept. of Language Science and Technology

tprakap@coli.uni-saarland.de

Abstract

Nowadays, the influence of the state-owned media in authoritarian countries is drastic. Hence, it is important to recognise government manipulation to effectively resist it. In this research, I analyse Belarusian pro-governmental media - "SB. Belarus' Segodnya". This paper presents a new corpus of texts, 8 years of political news (20K articles). To explore the intended framing, I compare this data with the independent Belarusian media covering the same period of time. I also review the major political myths of the sovereign Belarusian regime ruling from 1994 until the present. To confirm the salience of several narratives, I conduct unsupervised topic modelling (LDA). Finally, I describe differences in narratives before and after the disputed presidential elections in August 2020. Using means of both count-based and topic modelling methods, I confirm that the topic coverage landscape was transformed.

1 Introduction

Media control is one of the most important tasks for authoritarian regimes. Overt techniques include fake news, falsified information, and censorship, which in combination with agenda-setting (choosing *what* to cover) and framing (choosing *how* to cover) can shape public opinion in the desired way. Whereas a lot of research is devoted to the analysis of Russian (Herpen 2015; Field et al. 2018) and Chinese (King et al., 2017) government manipulations, I decided to focus on characterising the regime in Belarus, ruled by Alyaksandr Lukashenka¹ from 1994 until the present. Despite the small influence of Belarusian state media narratives in shaping the global information environment, I find it valuable to research which myths and ideas were developed by Belarusian authorities.

¹I stick to Belarusian transcription of names, however, in citations the Russian version can emerge (Lukashenko)

Also, I would like to measure computationally how pro-governmental media react to major political events like the contentious presidential campaign in 2020.

To do this, I will analyse the main printed newspaper in the country - "SB. Belarus' Segodnya" ("SB. Belarus Today", SB for short). It is owned by the Administration of the President and is widely known for propagating pro-governmental agenda. Employees of this media are constantly targeted by EU and US sanctions starting from 2006. Motivation statements for sanctions include accusations in support and justification of the repression of the democratic opposition and civil society. For example, Andrei Mukavozchyk, a well-known propagandist, entitles his [article](#) '*Becoming adult*', saying that Belarusian society has to resist democratic activists with all possible means. SB journalist proposes to start with informing security forces about opposition-minded persons to later detain them or even kill, because '*they are cancer threatening our society*'. To conclude, this printed press is an illustrative example of the pro-government media in Belarus.

Contributions of this work are manifold:

- First, this paper gives an overview of the Belarusian political situation, and highlights the roles of pro-governmental and independent media in shaping public opinion.
- Secondly, the paper presents a new corpus of political news from the web version of "SB. Belarus' Segodnya". This collection includes almost 20k articles from 2014 to 2022.
- Finally, this paper analyses the data using both simple count-based methods and more complex topic modelling techniques (LDA). It shows that agenda-setting and framing issues can be explored using simple means. All code and data are available on [GitHub](#).

Plan of the paper. After reviewing major political myths developed by the sovereign Belarusian regime in section 2, I explain the role of media in Belarus in Section 3 and proceed to corpus creation details in Section 4. I explain the topic modelling process in Section 5 and discuss obtained results in Section 6.

2 Main Narratives

Belarusian sovereign regime developed several stable narratives and myths (Shraibman, 2018; Shaput-ska, 2016). They actively use Soviet ideas ('NATO threat', 'Great Patriotic War', 'Union of Belarus and Russia'), nevertheless, there are plenty of new ones. For instance, promoting Lukashenka's personality, social stability, sovereignty, and 'everything but no war'. However, it is rather flexible. Belarusian ruling elites quickly adapt to transformations in the political and socio-economic fields. In the following subsections, the main ideas will be illustrated.

2.1 Stability and sovereignty

Current Belarusian state officials justify any social unfairness or low level of life (Belarus is one of the poorest European countries) as high costs of maintaining stability and resisting external enemies. According to their statements, the sovereign Belarusian state is following a unique path between East and West and it can explain any deprivations that Belarusian people are experiencing now (*'In 10 years we have shaped a modern state by our own labour, intellect, and hands, and we have taken our own path of development not imposed by anyone.'*, Lukashenka's public speech in 2007).

This narrative is also helping to protect the regime from the threat of protests. They employ propaganda to exploit a historical fear of social upheaval among Belarusians, as well to discredit the idea of protesting itself (*'How the protests against the ruling party threaten Poland'*). This technique is a common characteristic of authoritarian regimes: they claim they are not violating human rights or constraining the opposition but merely protecting the people and domestic stability.

2.2 Lukashenka personality - 'father of the nation'

The main public actor of the political mythogenesis is the Belarusian president, Aliaksandr Lukashenka. Other actors are auxiliary. Propaganda cultivates

an image of Lukashenka as an experienced, reliable leader who brought the country out of the chaos of the 1990s. Lukashenka himself likes to attribute even non-political accomplishments solely to his virtues (*'On my shoulders, they [Piesniary band] managed to preserve their existence through times of change'*, interview to Russian media in 2005). As a result, the cult of personality is rather functional than a pure cult of the leader. There are no streets and cities named after Lukashenka, no statues or large museums devoted to him.

2.3 Soviet heritage

Belarus is commonly recognized as one of the most Soviet-influenced states. There are a lot of streets named after communists (Satsukevich, 2021), Lukashenka treats USSR dissolution as a *global world catastrophe* and calls himself a *'soviet man'*. Hence, there is no surprise that Belarusian officials and state-owned media use the narratives from Soviet times (I take all citations from Lukashenka's public speech in 2022):

- treating collective West and NATO as threats; *'Military infrastructure is being strengthened in the West. Offensive in nature, by the way.'*
- cultivating Soviet myth of Great Patriotic War; *'In the national character is the spirit of the heroism of the ancestors who defended the country in the Great Patriotic War.'*
- highlighting the importance of the Union of Belarus and Russia to resist external enemies. *'If aggression is committed against our country, there will be hundreds of thousands of Russian troops here, who with hundreds of thousands of Belarusians will defend this sacred land.'*

3 Media in Belarus

In this section, I would like to characterise the media landscape in Belarus, and the shift in topic coverage after the presidential campaign in 2020. To start with, I distinguish two big blocks: state-owned and independent (including bloggers) media.

3.1 State-owned media

State-owned media are mostly represented by TV channels and printed media. According to the *official web page of the President of Belarus*, the largest national mass media outlets are Belarusian TV and Radio Company, the Second National TV

Channel (ONT), ZAO Stolichnoye Televideniye (STV). ONT is largely relying on Russian First Channel programs. The printed press is represented by the publishing house Belarus Segodnya and the publishing house Zvyazda. The Belarusian government also owns the news agency - Belarusian Telegraph Agency (BelTA). All these media are forcibly spread among budget organisations.

3.2 Independent media

On the other hand, there are plenty of independent media. The most popular ones are represented on the Internet because printed media and TV channels are heavily controlled by the government. I see the following categories of independent media: 1) Large media with Belarusian origin and financing: TUT.BY (now [Zerkalo.io](#), considered as the most popular media in Belarus), [Onliner.by](#), [Nasha Niva](#). 2) Media financed by US and EU: [Belsat](#), [Radio Svyaboda](#). 3) Bloggers (Telegram, YouTube): [Motolko](#), [NEXTA](#).

In highly controlled states Internet-based initiatives and bloggers are traditionally gain popularity. Belarus is not an exception ([Herasimenka, 2016](#)). Besides numerous blogging initiatives, the main media outlet in Belarus was an exclusively web project. [TUT.BY](#) was considered to be one of the most popular web pages in Belarus before its crack-down by the government in May 2021. According to the Gemius service, in April 2021, the portal covered 63% of the Belarusian internet audience. 3.3 million people or every third resident of Belarus visited TUT.BY. Their website traffic was 16 times higher than the website traffic of the largest Belarusian state news agency belta.by and is comparable to the traffic of popular Polish and Russian publications – [Gazeta.pl](#) and [Komsomolskaya Pravda](#) in Russia. TUT.BY still owns [a YouTube channel](#) with 1.45 million subscribers (March 2022).

Hence, I can conclude that despite the tight state regulation, Belarusians are active Internet users and they almost ignore state propaganda thanks to the alternative sources of information.

3.3 Presidential campaign in 2020

The presidential campaign in 2020 was the major political event in the recent history of Belarus. The low influence of traditional state-owned media and high level of Internet usage among Belarusians allowed people to organise themselves effectively to resist the authoritarian government. Officials used their traditional methods to intimidate civil soci-

ety and stop the spread of information. Popular bloggers Tsikhanouski and Losik were detained several months before the actual election date, licenses for foreign media were revoked and [their journalists were arrested for short sentence](#). However, all efforts of the ruling elites were not able to stop Belarusians from mass protesting in August 2020. Several hundred thousand people [gathered in Minsk on 16 August](#), a week after the election date. Later, Lukashenka's regime was able to crack down on a protest using massive support from Russia, including a media one. Specialists from Russia-Today were sent to Belarus in August 2020 to build a media narrative and hide the effect of massive strikes amongst former pro-government journalists ([Viačorka, 2021](#)). I will define and explore major challenges for state-owned media after the revolution later in Section 6.

4 Corpus Creation

To analyse pro-governmental media computationally, I have decided to create a corpus of political articles by "SB. Belarus' Segodnya" from 2014 to 2022. It is the largest newspaper in Belarus, owned by the Administration of the President. The weekly circulation is about 190k exemplars. The main source of the data was the [web version](#) of the newspaper.

4.1 Data Scraping

We have decided to collect exclusively political articles to focus on agenda-setting and framing analysis later, following the techniques described in [Field et al. \(2018\)](#). Online articles are available from 2014 until the present time. The algorithm was the following:

1) I started from the generic seed URL and iterated by aggregation pages simply changing page numbers from the oldest to the newest. On each aggregation page, around 20 links to the articles were scraped.

2) To store data I utilised a local [SQLite](#) database. This data organisation method allows usage of SQL queries for interactive analysis (current document count, finding duplicates, selecting articles from date range and looking for particular substrings in document body). Articles are stored with the following structure (I will illustrate each field with an example from actual data):

- **document_id** - non-empty unique identifier of the document: *1440895*

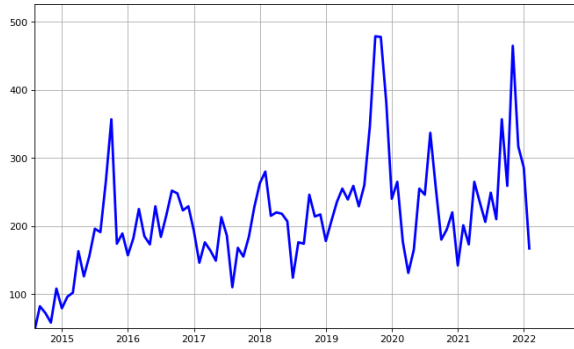


Figure 1: Number of articles per month. Peaks are corresponding to important political events.

- **title** - primary document title: *To whom and why does union integration hinder?*
- **title.h1** - optional secondary document title: *Our alternative*
- **tags** - optional list of tags selected by author: *union state@politics@economy@sanctions*
- **similar_documents** - optional links to similar articles: <https://www.sb.by/articles/soyuz-belarus-rossiya/>
- **publication_date** - datetime of publication: *2021-10-21T10:42:00+03:00*
- **author** - name of the author, which I lower-cased during parsing: *polina konoga*
- **body** - main text of the article: *Alexei Avdonin: The West has chosen the Union State as its target.* (first sentence, full body is much longer)
- **hyperlink** - original url of the article: <https://www.sb.by/articles/nasha-alternativa-souz.html>

There were several issues during parsing and scraping. Some of them were technical: http/https references to the same pages; varying data organisation inside pages. Other problems were content-related: old documents do not contain tags and references to similar documents; some pages are not articles, but special pdf editions of the paper or links to external sources. All these issues were addressed using a set of simple heuristics.

Finally, a corpus of 19334 documents was created.

4.2 Initial Data Analysis

In this section, I conduct an initial analysis of the collected corpus. Even simple methods, based on words or tags count, can reveal intrinsic media manipulation strategies like agenda-setting. Addition-

Full Corpus	'Before'	'After'
Lukashenko	Lukashenko	Lukashenko
Belarus	President	Belarus
President	elections	MID
MID	USA	Russia
Russia	Russia	Makey
USA	MID	opinion
elections	Ukraine	ODKB
Makey	Belarus	Poland
Ukraine	Makey	refugees
Putin	Trump	USA

Table 1: Top-10 Tags. Full corpus contains data from 2014 until 2022; 'Before' corpus - 2014 till May 2020; 'After' corpus - May 2020 till 2022.

ally, major political events can be identified by looking at the number of articles for the specific periods. In the following subsections, I will describe the preliminary corpus analysis in detail.

4.2.1 Tags Analysis

Almost every article starting in 2016 has special tags, selected by the author. Count-based analysis can help to identify the most popular topics. Full lists of top-10 popular tags are provided in Table 1. I consider the full dataset (2014-2022) and two subsets: from 2014 till May 2020 ('Before') and from May 2020 till 2022 ('After') to compare tags usage shift after the presidential campaign in 2020. I keep some original abbreviations (MID - Ministry of Foreign Affairs, ODKB - Collective Security Treaty Organization) while translating others (USA).

Speaking about the full corpus, the most popular tag is 'Lukashenko' (3432), the second place takes 'Belarus' (1020), and the third one is 'President' (832). Other popular tags include 'Russia', 'USA', 'Ukraine', 'elections'. Interestingly, SB journalists stopped active use of the tag 'President' after unfair elections. Instead, the usage of the tag 'Belarus' increased. Also, the paper introduced the 'opinion' tag to mark a new type of article with the personal opinion of pro-governmental propagandists.

4.2.2 Articles per month

The second direction of preliminary analysis is counting the number of articles per month (Figure 1). The mean value of publications per month is 212.5, but there are several peaks (above 300 articles) allowing detection of important political events in Belarus. For example, the highest peak

Topic Name	Topic words extracted by LDA modelling
Lukashenko	Lukashenko, president, tell, know, want, understand
Europe	geopolitical, European, dialogue, threat, West, confrontation
Soviet history	GDR, monument, Czechoslovakia, Stalin, Gorbachev, socialism
Union with Russia	union, Belarusian-Russian, eurasian, Putin, integration
Migrant Crisis	refugee, Polish, migrant, Lithuania, police
Military	ODKB, drills, Zas', defence, Afganistan, army
Internal European Politics	Macron, Catalonia, Merkel, Brexit, Scotland, eurosceptic
Economy	market, budget, export, bank, financial, ruble
Belarusian Elections	candidate, voting, elections, Ermoshina, observer
Belarusian Words	what, about, this, country, also, Aliaksandr, Belarusian

Table 2: Selected topic words from full corpus, representing articles from 2014 till 2022.

is detected at the end of 2019. This period included parliamentary elections, pressure from Russia about closer integration, and the following oil supply crisis. Also, there are peaks corresponding to the presidential elections (October 2015, August 2020) and the migrant crisis (November 2021). Presidential elections in 2015 and parliamentary elections in 2019 were covered more than a very active political campaign in the summer of 2020. In my opinion, it can be explained by intended agenda-setting.

4.2.3 Count-based word analysis

To perform word analysis, I did the preprocessing including the following steps: lowercasing, tokenization, lemmatization, excluding stopwords and punctuation signs. Lemmatization was performed using the MyStem tool (Segalovich, 2003). After preprocessing there are 4,764,694 (111,786 unique) tokens in the corpus.

The last step is a count-based analysis of the most popular tokens. For the full corpus, the most frequent tokens are 'Belarus', 'president', and 'country'. The set of most popular tokens before the presidential campaign in 2020 resembles the tokens set of the full corpus, whereas a different picture is observed after the campaign. 'Belarus' is still the most popular token, but now 'country', 'our', 'state' occupy positions in the top-4 leaving 'president' token only at 5th place.

Detailed analysis of the above-mentioned findings in terms of framing and agenda-setting will be performed later, in Section 6.

5 Topic Modelling

Starting with simple count-based methods I proceed to more complex analysis. Latent Dirichlet

Allocation (Blei et al., 2003) was used to conduct topic modelling. It is a generative statistical model that allows sets of observations to be explained by unobserved groups. I followed a tutorial by Darling (2011) to implement the LDA algorithm. Then, I run it with the following hyperparameters: 300 iterations, 20 topics, 0.02 alpha, and 0.1 beta. To interpret the results, I used the LDAvis library by Sievert and Shirley (2014). The relevance parameter was set to 0.2 to filter out the most frequent words. Topic names in the further analysis were inferred from identified word lists, so this analysis is subjective to some extent.

5.1 Full Corpus

First, a full corpus is examined. Details are provided in Table 2. Interestingly, Belarusian, Polish, and other foreign words (English and Spanish) formed three separate lists, assembling words without strong semantic connotation as for other topics.

Further topics are more connected semantically. The first and the biggest topic is devoted to Lukashenko's personality and his actions illustrated by the following verbs: 'tell', 'know', 'want', 'understand'. In my opinion, it accentuates the key importance of the president and his actions in Belarus.

Soviet history, especially the part about large influence in Europe (GDR, Czech Republic) is confirmed as an important aspect of Belarusian sovereign political mythology.

We see a separate topic about Union with Russia, thereby confirming the salience of this myth. LDA also discovered a set of words representing a military operations topic. The most illustrative words here are 'defence', 'drills', and 'ODKB' (Collective Security Threat Organisation, Russia-headed

Topic Name	Topic words extracted by LDA modelling
Religion, Sport and History	olympic, memory, church, deed, monument
Ukraine Crisis	Poroshenko, Zelenskij, Donbass, contact, Kyiv, Normandy
Migrant Crisis	migrant, refugee, Hungary, Czechia, Germany, Polish

Table 3: Selected topic words from 'Before' corpus, representing articles from 2014 till May 2020.

Topic Name	Topic words extracted by LDA modelling
Western Sanctions	sanction, West, pressure, action, interference
Fascism and Church	memory, nazism, heroic, orthodox, patriotic, upbringing
China Partnership	China, industrial, partnership, strategic, accomplishment, global
Parliament and Laws	assembly, house, parliamentary, commission, deputy, legal
Protests	police, demonstrator, protest, hurt, investigator, detainee

Table 4: Selected topic words from 'After' corpus, representing articles from May 2020 till 2022.

military union). Afghanistan is mentioned as a potential source of the conflict, where ODKB should potentially react and protect its member state, Tajikistan.

Affairs with Europe are described with contradictory terms like 'conflict' and 'dialogue', although negative sentiment is prevalent ('threat', 'confrontation'). Balancing between Russia and other allies is one of the main characteristics of Belarusian foreign politics. Nevertheless, the conclusion can be drawn that Russia is considered an ally, and the West an enemy.

Speaking about temporary topics, I see the one devoted solely to the migrant crisis. This topic is mostly about the recent conflict on the Polish and Lithuanian borders in November 2021. There is an evidence that this small topic was covered extensively because it is distinguishable even in the full corpus with articles from 2014 to 2022. Another temporary topic, which I called 'Internal European Politics', includes several eurosceptic processes like Brexit or referendums in Scotland and Catalonia.

Some neutral topics, like the economy or Belarusian elections, were also discovered.

5.2 Before Presidential Campaign 2020

Details are provided in Table 3. Here I will focus on topics that are not present in the full corpus because it is obvious that the 'before' corpus has a large presence in the full version of data.

There is a topic 'Religion, Sport and History', which combines very different sub-topics. In my opinion, it can be defined as the ideological basis of the sovereign Belarusian regime.

We also observe some topics, that can be de-

scribed as adaptive or temporary. For instance, SB actively covers the crisis in Ukraine that started in 2014. Also, there is a clear topic about the migrant crisis in Europe in 2015.

5.3 After Presidential Campaign 2020

Details are provided in Table 4. As expected, there is a clear anti-West topic. It is mostly about sanctions, pressure, and interference in sovereign Belarusian affairs.

Crucial partnership with China is heavily highlighted. This country is depicted as a global leader, that actively supports Belarus ('global', 'strategic', 'partnership'). It is worth mentioning, that allying with China was described earlier in SB topics, discovered by LDA, but it was mixed with other countries.

For the first time, there is a single topic about parliament and laws. Previously, it was mixed with the judiciary. To my mind, it can be explained by the increased role of parliament in passing various anti-protest legal acts. After careful manual examination of this topic in the original data, I see a lot of interviews with members of both parts of the Belarusian parliament. From my standpoint, it could be explained by the desire to create a feeling of mass support for the government's actions, since deputies are supposed to be elected by the people (which is not actually true in Belarus).

Finally, I found a topic about protests and police repressions.

6 Discussion

Here I would like to link various data-driven phenomena, discovered in previous sections, with con-

cepts of framing, agenda-setting, and topic modelling.

6.1 Agenda-Setting

To detect agenda-setting, I have decided to compare SB data with independent media (TUT.BY) coverage for the same period.

Sinyavski (2021) provides topic models built on TUT.BY articles as well as source data from this independent media. I will be analysing the frame from 15.05.2020, the beginning of the election campaign, to 11.05.2020, the latest available data from TUT.BY. For this period SB has 2620 political articles compared to 41455 various articles from TUT.BY (all available data was scraped). I utilise SQL queries for surnames matching in the text of the articles.

SB journalists mention Sviatlana Tsikhanouskaya, the leader of the Belarusian democratic movement, and her husband, political activist Siarhei Tsikhanouski, (I was looking for 'Tsikhanousk' substring) only 78 times. Another popular independent candidate Viktor Babaryka was mentioned 15 times, whereas Aliaksandr Lukashenka has 1075 (41% of all articles) entries.

I used the same SQL query to check articles about the Tsikhanouski family in TUT.BY materials and found 1733 entries. There are 1176 articles about popular opposition candidate Viktor Babaryka, who was detained in June 2020. Sure, TUT.BY was a larger media than SB, but the difference in mentions rate is around 120 times (15 vs. 1176). For the sake of complete comparison, Lukashenka was mentioned 4335 (10% of all articles) times.

Hence, there are clear signals of agenda-setting (a much larger proportion of articles about Lukashenka in SB materials, downplaying the role of Babaryka and Tsikhanouski family in the election campaign), detected by the comparison.

6.2 Framing

To identify framing issues, I closely examined 15 articles about Viktor Babaryka (a popular opposition candidate, detained a month after the start of the campaign) in SB. One of them claims that Babaryka is a 'deceiver' and 'definitely not law-abiding businessman'. Another one cites Lukashenka's speech, where he invites Interpol experts to Belarus to make sure that detention of Babaryka was lawful and non-political.

On the other side, TUT.BY covers this topic in much more detail (374 articles mentioning the Babaryka case). Independent journalists raise concerns about the lawfulness of propagandistic stories, depicting Babaryka as a criminal before the court decision. TUT.BY also mentions appeals of defense lawyers to Prosecutor General of Belarus and Minsk City court, thereby showing readers that this case is not so clear.

Framing can be observed as well in the list of most-popular tokens in SB articles after May 2020 ('Belarus', 'country', 'our', 'state'). To identify and classify the issue in words from the new top, I refer to 'The Policy Frames Cookbook'. Boydston et al. (2014) present 15 cross-cutting framing dimensions, such as economics, morality, and politics. This research was informed by the framing literature and developed to be general enough to be applied to any policy issue. To be specific, the 'Cultural identity' frame includes traditions, customs, or values of a social group in relation to a policy issue. Words from the new top ('country', 'our', 'state') are fitting to this frame.

6.3 Topic Modelling Analysis

In this subsection, I would like to focus on newly emerged narratives in Belarusian state-owned media after the presidential campaign in 2020 and confirm them by LDA topics.

6.3.1 Explain the revolution

Following narratives of state-owned Belarusian media, Lukashenka has enormous support. According to official election results, he obtained more than 80% of the votes in 2020. Thereby, the major tasks of government officials and state-owned media included an explanation of how mass protests were possible at all.

Ruling elites started to promote a myth of enemies, accusing the collective West (EU and US) of financing protests and interfering to internal Belarusian affairs. Speaking with BBC journalist in 2021, Lukashenka claimed that they are destroying all non-government and non-profit organisations, because the collective West was financing revolution through them.

The anti-West narrative is observed in LDA topics ('sanction', 'West', 'pressure') as well. However, it is mostly about sanctions pressure and interference, not about protests financing. To sum up, a general myth of foreign enemies, having roots in the Soviet Union, got a new life.

6.3.2 Diminish protesters

The second main aim, which I draw from the public government's discourse, was to diminish the protesters themselves. For this goal, state-owned media used several narratives to show that only marginalised persons are supporting the revolution.

First, they claimed that the white-red-white flag, that extensively used by protesters, is tied with Nazi collaborators during Second World War. Also, state-owned media and Belarusian officials started to promote another political myth actively: the historical memory of the Great Patriotic War. They were trying to convince people that the ongoing repressions are targeting not civil society, but neo-fascists.

Speaking about LDA-based analysis, there is a topic mixing words from two sub-narratives: fascism and church. Journalists of "SB. Belarus' Segodnya" are trying to link protesters with nazis, enemies of Belarus during the Great Patriotic War. They appeal to historic memory, heroic deeds, and promoting unity to resist enemies. The role of the Orthodox Church here is to support a government in a patriotic upbringing of the citizens.

Second, state-owned media used directed hate speech. For example, Ryhor Azaronak, one of the well-known propagandists, [calls](#) activists of the anti-regime movement "inhuman", "parasites and idlers", "rats". Dehumanization is often accompanied by death threats. Andrei Mukavozchyk, another widely recognised propagandist, [writes](#): "We will definitely find you ... and we will hang you, side-by-side," about two UK dissidents of Belarusian origin.

6.3.3 Hide repressions

The last but not least challenge for Lukashenka's regime in Belarus was to hide brutal repressions and their real scale.

Firstly, Belarusian officials use state-owned media to falsify information. Akrestina, a notorious detention center in Minsk, is known as a place where thousands of people were tortured in August 2020 and later. Trying to deny this information, Ryhor Azaronak [filmed a video](#), picturing Akrestina as a safe and clean place.

Secondly, the regime destroys major independent media, which continue honestly describing the situation. TUT.BY web-page was [blocked in May 2021](#), and 15 employees were imprisoned. Yavor Martsinovich, editor-in-chief of the oldest Belarusian newspaper - Nasha Niva - was [detained](#)

[in July 2021](#). Moreover, opposition media activist Raman Pratasevich was arrested after the [forced landing of Ryanair plane](#) in May 2021.

LDA assisted in detecting this intent as well, highlighting the topic of protests (Table 4: 'police', 'demonstrator', 'detainee'). A closer analysis of this narrative was performed using SQL query looking for the 'demonstrator' sub-string in the documents' bodies. This search revealed that journalists of "SB. Belarus' Segodnya" mostly write about upheavals in Kazakhstan, Kyrgyzstan, or covid-related protests in the EU and US. In my opinion, it proves that the topic of Belarusian protests is intentionally uncovered.

6.4 Future Work

I see the following directions for future work. First, other topic modelling techniques can be utilised. For example, top2Vec ([Angelov, 2020](#)) uses contextual word embeddings, and does not require a predefined number of topics and extensive pre-processing. Second, I am going to create a corpus of materials written by recognised propagandists, like Andrei Mukavozchyk. Further analysis of this data can include hate speech detection and shift in narratives after the presidential elections in Belarus in 2020. Third, word co-occurrence can be examined more properly for framing detection. I expect that mostly negative adjectives are employed to describe the opposition, whereas the actions of ruling elites are shown as extremely wise and error-free.

7 Conclusion

In this paper, I described political myths created by Lukashenka's regime in Belarus and the role of propaganda in it. I crafted a corpus of 20k political articles from the main pro-governmental printed press in Belarus - 'SB. Belarus' Segodnya'. To analyse collected data, both count-based and statistical (LDA) documents analysis were performed. I reviewed in detail the shift in narratives of state-owned media news coverage after the rigged presidential elections in 2020 and the following protests. Computational means assisted in the confirmation of certain strong narratives in the selected state-owned media. Additionally, I revealed cases of agenda-setting and framing by comparing this corpus with Belarusian independent media covering the same period.

References

- Dimo Angelov. 2020. [Top2vec: Distributed representations of topics](#).
- David Blei, Andrew Ng, and Michael Jordan. 2003. [Latent dirichlet allocation](#). *Journal of Machine Learning Research*, 3:993–1022.
- Amber E. Boydstun, Dallas Card, Justin H. Gross, Paul Resnick, and Noah A. Smith. 2014. [Tracking the development of media frames within and across policy issues](#). *APSA 2014 Annual Meeting Paper*.
- William M. Darling. 2011. [A theoretical and practical implementation tutorial on topic modeling and gibbs sampling](#).
- Anjalie Field, Doron Kliger, Shuly Wintner, Jennifer Pan, Dan Jurafsky, and Yulia Tsvetkov. 2018. [Framing and agenda-setting in Russian news: a computational analysis of intricate political strategies](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3570–3580, Brussels, Belgium. Association for Computational Linguistics.
- Aliaksandr Herasimenka. 2016. [Political participation and internet platforms: How new communication technologies help belarusian civic activists](#). *Politics and Society in Belarus*, 1:4–12.
- Marcel H Van Herpen. 2015. *Putin’s Propaganda Machine: Soft Power and Russian Foreign Policy*. Rowman & Littlefield.
- Gary King, Jenniefer Pan, and Margaret E. Roberts. 2017. [How the chinese government fabricates social media posts for strategic distraction, not engaged argument](#). *American Political Science Review*, 111(3):484–501.
- Ivan Satsukevich. 2021. [The soviet heritage in street names of belarus](#). *Communist Heritage in Belarus and EU Countries: The Problem of Interpretation and the Relevance of Conservation*, 1:190–199.
- Ilya Segalovich. 2003. [A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine](#). In *Proceedings of the International Conference on Machine Learning; Models, Technologies and Applications*, pages 273–280.
- Tatsiana Shaputska. 2016. [Political mythogenesis in belarus in 1994-2010: ruling elite’s discourse](#). *Politics and Society in Belarus*, 1:13–25.
- Artyom Shraibman. 2018. *The house that Lukashenko built: the foundation, evolution, and future of the Belarusian regime*. Carnegie Endowment for International Peace, Washington DC.
- Carson Sievert and Kenneth Shirley. 2014. [LDAvis: A method for visualizing and interpreting topics](#). In *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pages 63–70, Baltimore, Maryland, USA. Association for Computational Linguistics.
- Timur Sinyavski. 2021. [Construction of semantic vector representations of text documents](#). *Belarusian State University*. In Russian.
- Franak Viačorka. 2021. *The infowar behind the Belarus revolution*. Atlantic Council, Washington DC.