

Vol. 11 No. 1 (2025)

# intersections

East European Journal of Society and Politics

**Text as Data**

Eastern and Central European political discourses from  
the perspective of computational social science – Part 2

EDITOR-IN-CHIEF:  
Attila Bartha

EDITORS:  
Jekatyerina Dunajeva, Margit Feischmidt, Ákos Huszár,  
Nóra Kovács, Miklós Könczöl, Eszter Neumann,  
Orsolya Ring, Bence Ságvári, Veronika Szontagh, Márton Varju

MANAGING EDITOR:  
Arin Agich

EDITORIAL ASSISTANT:  
Andrea Demeter

COPY EDITORS:  
Simon Milton  
Chris Swart

Guest editors of the thematic issue:  
Domonkos Sik, Renáta Németh, Ildikó Barna, Theresa Gessler & Hanna Orsolya Vincze

PUBLISHED BY:



ELTE Centre for Social Sciences  
Zsolt Boda  
Director General

<https://intersections.tk.hu>  
E-ISSN: 2416-089X

Intersections. East European Journal of Society and Politics is an Open Access, double blind peer-reviewed online journal. When citing an article, please use the article's DOI identifier.

The publication of *Intersections. East European Journal of Society and Politics (IEEJSP)* was supported by HUN-REN Hungarian Research Network between 1 January, 2024 - 31 July, 2025.

---

## Table of Contents

---

### EDITORIAL

DOMONKOS SIK, RENÁTA NÉMETH, ILDIKÓ BARNÁ, THERESA GESSLER & HANNA ORSOLYA VINCZE	
Preface to the special issue ‘Text as data’ – Eastern and Central European political discourses from the perspective of computational social science (Part 2)	1

### RESEARCH PAPERS

RENÁTA NÉMETH, ESZTER KATONA, PÉTER BALOGH, ZSÓFIA RAKOVICS & ANNA UNGER	
What else comes with a geographical concept beyond geography? The renaissance of the term ‘Carpathian Basin’ in the Hungarian Parliament	3
EMESE TÚRY-ANGYAL & LÁSZLÓ LŐRINCZ	
Sharing Political News Online: A Network Model of the Spread of Information on Social Media	40
RADU M. MEZA & ANDREEA MOGOS	
Fear and Loathing in Eastern Europe. Framing and Agenda Setting of Foreign Affairs in Czech, Polish, and Romanian Sputnik News (2017–2022) – An Analysis of Headlines	63
ILYA SULZHYTSKI, & VARVARA KULHAYEVA	
Ambassadors of War: Social and Semantic Networks of Belarusian Pro-Government Telegram Channels during the 2022 Russian Invasion of Ukraine	86
ROK SMRDELJ, ROMAN KUCHAR & MONIKA KALIN GOLOB	
Expanding Boundaries: “Gender Theory” and the Twitter (X) Debate on Gender-Sensitive Language Use in Slovenia	111
KATA KNAUZ, ATTILA VARGA, ZSOLT SZABÓ & SÁRA BIGAZZI	
Mainstream Political Discourse on the Roma Minority in Hungary between 2010 and 2019	139
ARON SZALAY & ZSÓFIA RAKOVICS	
Tuned to Fear – Analyzing Viktor Orbán’s State of the Nation Addresses, focusing on the enemy images identified in the National Consultation	165

---

---

ÁRPÁD RAB, TAMÁS SZIKORA, & BERNÁT TÖRÖK

Towards a trustful Information Society.

Comparative analysis of the Czech Republic,  
Hungary, Poland and Romania

185

## **BOOK REVIEW**

Singh, J. (2023). Natural Language Processing in the Real World: Text Processing,  
Analytics, and Classification. CRC Press, Taylor & Francis Group

Reviewed by MOHAMMAD ASHRAFUL ALAM

192

---



---

DOMONKOS SIK,\* RENÁTA NÉMETH,\*\* ILDIKÓ BARNA,\*\*\*  
THERESA GESSLER\*\*\*\* & HANNA ORSOLYA VINCZE\*\*\*\*\*

Intersections. EEJSP

11(1): 1-2.

<https://doi.org/10.17356/ieejsp.v11i1.1471>

<https://intersections.tk.hu>

## Preface to the special issue 'Text as data' – Eastern and Central European political discourses from the perspective of computational social science (Part 2)

- 
- \* [\[sik.domonkos@tatk.elte.hu\]](mailto:sik.domonkos@tatk.elte.hu) (Eötvös Loránd University, Faculty of Social Sciences, Budapest)  
\*\* [\[nemeth.renata@tatk.elte.hu\]](mailto:nemeth.renata@tatk.elte.hu) (Eötvös Loránd University, Faculty of Social Sciences, Budapest)  
\*\*\* [\[barna.ildiko@tatk.elte.hu\]](mailto:barna.ildiko@tatk.elte.hu) (Eötvös Loránd University, Faculty of Social Sciences, Budapest)  
\*\*\*\* [\[gessler@europa-uni.de\]](mailto:gessler@europa-uni.de) (European University Viadrina, Frankfurt/Oder)  
\*\*\*\*\* [\[vincze.orsolya@fspac.ro\]](mailto:vincze.orsolya@fspac.ro) (Babeş-Bolyai University, Cluj-Napoca)

This issue is the second part of the 'Text as data' – *Eastern and Central European political discourses from the perspective of computational social science* special issue, the first part having been published as the issue 2024/4. The articles all approach social research problems through text analysis, using computational techniques, typically natural language processing (NLP).

In the second part of the special issue, the Central European political discourses are mapped from different angles. Renáta Németh, Eszter Katona, Péter Balogh, Zsófia Rakovics and Anna Unger explore the discourses surrounding the Carpathian Basin, a central metaphor of a collective identity anchored in the narratives of historical Hungary from parliamentary debates. Emese Túry-Angyal and László Lőrincz analyse discursive mechanisms, partly related to the technological infrastructure of algorithmic public spheres: they explore how echo chambers, homophily and network type affect the spread of information on Facebook.

According to Radu M. Meza and Andreea Mogoş generating fear and loathing is not a rare tendency in Central European political discourses. By analysing the headlines of Sputnik News (broadcast in Poland, Romania and the Czech Republic), they mapped the discursive strategies, and the affective framing that appear in the communication of a Russian government-funded news agency. In a similar vein, Ilya Sulzhytski and Varvara Kulhayeva analyse Belorussian Telegram channels from the perspective of their interpretation of the Ukrainian invasion. Besides mapping the main discursive panels, they also demonstrate that local pro-government activists are central channels of the creation and dissemination of hatred towards Ukrainians in Belarus. Although the impact of war is the most tangible in the present, the collective traumas also linger in collective memories.

Diagnoses of the vulnerability of Central European democracies can be refined by reflecting on the ongoing discursive mechanisms. One of the most dangerous potentials of political polarization and the consequent intensification of social conflicts are indicated by the various forms of scapegoating. Rok Smrdelj, Roman Kuhar and Monika Kalin Golob analyse another hotspot in contemporary identity politics, namely the debates surrounding gender. Based on Twitter posts, they attempt to map filter bubbles of anti-gender

---

discourses. Kata Knauz, Attila Varga, Zsolt Szabó and Sára Bigazzi analyse the Hungarian political discourses of the Roma minority. Besides the explicit stereotypes and hostile prejudices, an attempt is also made to detect the more subtly biased, paternalistic discourses, which appear in the communication of parties. Áron Szalay and Zsófia Rakovics analyse the enemy images appearing in the speeches of Viktor Orbán, the prime minister of Hungary. Besides mapping the range of generic (e.g. migrants) and specific (e.g. Soros, Brussels) enemies, the discourses of fearmongering are also examined in detail.

The special issue concludes with two shorter studies. In his book review, Mohammad Ashraful Alam writes about the book *Natural Language Processing in the Real World: Text Processing, Analytics, and Classification*, written by Jyotika Singh, which is a practical guide for building natural language processing solutions. The special issue also contains a short data note on a comparative analysis of information society in Central and Eastern Europe by Árpád Rab, Tamás Szikora and Bernát Török. Findings reveal both regional commonalities and distinct national attitudes towards online manipulation, social media use and the impact of internet communication on personal relationships.

---

RENÁTA NÉMETH,\* ESZTER KATONA,\*\* PÉTER BALOGH,\*\*\*  
ZSÓFIA RAKOVICS\*\*\*\* ✉ ANNA UNGER\*\*\*\*\*

Intersections. EEJSP  
11(1): 3–39.  
<https://doi.org/10.17356/ieejsp.v11i1.1241>  
<https://intersections.tk.hu>

## What else comes with a geographical concept beyond geography? The renaissance of the term 'Carpathian Basin' in the Hungarian Parliament

---

- \* [\[nemeth.renata@tatk.elte.hu\]](mailto:nemeth.renata@tatk.elte.hu) (ELTE Eötvös Loránd University, Faculty of Social Sciences, Research Center for Computational Social Science)  
\*\* [\[katona.eszter@tatk.elte.hu\]](mailto:katona.eszter@tatk.elte.hu) (ELTE Eötvös Loránd University, Faculty of Social Sciences, Research Center for Computational Social Science)  
\*\*\* [\[peter.balogh@ttk.elte.hu\]](mailto:peter.balogh@ttk.elte.hu) (ELTE Eötvös Loránd University, Institute of Geography and Earth Sciences, Department of Social and Economic Geography; HUN-REN Centre for Economic and Regional Studies, Institute for Regional Studies, Transdanubian Research Department)  
\*\*\*\* [\[zsofia.rakovics@tatk.elte.hu\]](mailto:zsofia.rakovics@tatk.elte.hu) (ELTE Eötvös Loránd University, Faculty of Social Sciences, Research Center for Computational Social Science,)  
\*\*\*\*\* [\[unger.anna@tatk.elte.hu\]](mailto:unger.anna@tatk.elte.hu) (ELTE Eötvös Loránd University, Faculty of Social Sciences, Research Center for Computational Social Science)

### Abstract

A key tenet from research on geographical concepts is that these are never neutral but filled with different ideas and agendas. The 'Carpathian Basin' is one of the most significant concepts in Hungarian geographical thought, but its recently reemerging use in political discourse has not yet been studied through quantitative text analysis.

In this paper, we describe how a structural topic model was used to analyze the 1,525 speeches containing the term delivered in the Hungarian Parliament between 1998 and 2020.

Our results indicate a renaissance in the use of the term, both in terms of its more frequent use and its discursive meaning as a sign of a turn in national policy. At the same time, 'Carpathian Basin' discourse serves as a symbolic battleground for different political ideologies to indicate both neutral geographical references and nationalist sentiments. Left-liberals tend to use it politically neutrally, referring to an ethno-culturally heterogeneous area, and using a less personal voice, referring to institutions and interests. In contrast, right-wing narratives often demarcate the Carpathian Basin as a single geographical entity. Some of these speeches exhibit virtual nationalism, while others subtly question territorial legitimacy. The latter MPs speak in terms of representing their own community, referring to values, emotions, and culture, offering a collective identity to which people attach values and emotions.

**Keywords:** natural language processing, structural topic model, Carpathian Basin, Hungarian Parliament, ideological divides

---

## 1 Introduction

Language and concepts influence our thoughts and actions, and this is also true of geographical names. However, while, for instance, country names are well-established, spaces that lack formal institutions (including borders) need to be denominated in order to be imagined as real. The respective names then need to be continuously invoked and re-invoked if the intention is to build a consensus around the existence of that space.

Throughout various periods, the Carpathian Basin has been one of the most significant concepts in Hungarian geographical thought, including serving irredentist goals in the interwar era (Balogh, 2021). While the irredentist element has never been part of official policy since WWII, the concept has been gradually revived in recent decades (Hajdú, 2018) and now forms part of everyday political discourse and national identity construction (Scott & Hajdú, 2022). This article investigates how the concept has recently been used by Hungarian political elites. By examining speeches delivered in the Hungarian Parliament between 1998 and 2020 that contain the label ‘Carpathian Basin,’ we seek to answer the question of what latent themes can be distinguished in the discourses tied to the term. Further, can politically motivated identity-making patterns be detected? How does the framing of each political-ideological bloc differ? What changes can be observed during the studied period? Since the issue of national identity played a very important role in political debates during the period under examination, with a widening gap between political blocs regarding the concept of the nation, it can be assumed that these processes are also reflected in discourses related to the Carpathian Basin.

To examine the texts, we used a natural language processing tool called structural topic modeling (STM) (Roberts et al., 2019). This model is suitable for exploring the latent topics of a corpus and has the advantage of being able to incorporate meta-variables, the role of which can be twofold: influencing either the frequency of the occurrence of topics (prevalence variable) or their framing (content variable). Thus, by representing political position and time as meta-variables, the method can help answer the following question: What are the nature and dynamics of the relationship between a political position and the framing of latent topics? The paper aims to contribute to the current academic discourse in two ways.

First, from a substantive perspective, with a comprehensive quantitative analysis of the Hungarian parliamentary discourse. The ‘Carpathian Basin’ is one of the most significant concepts in Hungarian geographical thought, but its recent reappearance in political discourse has not yet been studied in relation to Parliament. Second, from a methodological perspective, we demonstrate the wide applicability of quantitative methods in textual analysis. Prior studies on national identity or critical geography usually use a qualitative approach, but the approach we use provides quantitative evidence, and because we use a complete corpus, our approach cannot be criticized using the arguments often raised against qualitative methods (incomplete empirical base or selective source selection). The method we use (the structural topic model) is also not yet widely used in the field and, to our knowledge, has not been applied to Hungarian social data research, so we also hope that our analysis will inspire others.

The article is structured as follows. This introduction is followed by a theoretical chapter that lays out our conceptual approach. The subsequent chapter briefly introduces the historical background and current context of the use of the term ‘Carpathian Basin.’

The next chapter presents our data and the methodological framework underlying the study. In this framework, structural topic modeling is combined with a qualitative approach. The subsequent chapter is devoted to the results. Both quantitative and qualitative approaches are used to select the most appropriate model and interpret the results. The final chapter concludes the paper by summarizing the key findings and putting the results in a broader context. We will argue that while the 'Carpathian Basin' has been used across political blocs, it has been particularly embraced by the right. Moreover, differences can be detected between the political camps concerning how the concept is invoked.

## 2 Text as the empirical base

### 2.1 'Text as data' in social research

Our empirical research draws inspiration from the recent 'text as data' approach in social research and the related methodology for the automated mining of large text databases, namely natural language processing (NLP). One antecedent of this empirical social research trend is the narrative turn in the social sciences (Goodson & Gill, 2011), which has primarily sought to provide a methodological alternative to positivist research, seeing in the text the possibility of a self-reflexive approach to social phenomena. Another, not unrelated, discipline-specific antecedent is the science of language and politics, where language is seen as an indispensable tool for political action (Wodak & Forchtner, 2017; Müller, 2008). We should also mention research on the discursive construction of national history (Wodak, 2010) and national identity (Wodak et al., 2009).

### 2.2 Political agendas associated with geographical labels

While Foucauldian and other scholars have long established that language – including word choice – influences thought as well as agency, geographical names and discourses have only more recently become subjects of investigation from this perspective (Medby & Thornton, 2023; Müller, 2008). The field of critical geopolitics has greatly contributed to this, showing that geographical labels are never neutral but filled with various ideas and intentions (Dodds, 2019). In addition, spatial concepts can play a significant role in identity politics and foreign policy visions (Balogh, 2022), as well as collective identity-making (Paasi, 2016; Egry, 2020), which is particularly true when competing ideas exist concerning what and where the nation is and where it should be heading. Hence, the intentions and effects of meaning-making in relation to geographical concepts are of primary interest.

According to Lacoste, the nation is 'the fundamental geopolitical concept' (1997: 38) in the sense that it 'refers [...] fundamentally to language and territory' (ibid. 36). Consequently, this is the spatial entity into which most people are primarily socialized through various processes, one of which is the discursive dimension (Paasi, 2016). However, macro-regional constructs can also be the subjects of political or collective desire, not least when they have existed in a real or imagined glorious past. As Bassin (2012: 553) has described, such meaning-making often 'involves the construction of idealized pictures of national glory lost at some point in the remote past, and then the projection of this picture

as an aspirational vision for the future.’ It goes without saying that macroregion-building projects are often contested and even dangerous, although this depends on whether any given nationalism is of the ‘hot’ or the ‘banal’ sort (cf. Paasi, 2016).

### 3 The concept of the Carpathian Basin

The Carpathian Basin (Kárpát-medence) is the designation of a physical geographical entity in Central Europe. Although the term exists in English and other languages, it is a genuinely Hungarian one in the sense that – except in the natural sciences – it is rarely used (or even known) by non-Hungarians (Fejes, 2011). This is because the label defines a space in which Hungarians constitute the single largest ethnic group. Accordingly, the concept has also been and remains contested, especially among Hungary’s neighbors, but also by some domestic commentators (Balogh, 2021; Scott & Hajdú, 2022). Even in Hungary, the term ‘Carpathian Basin’ only appeared in the late 19th century, and even then, very sporadically (Balogh, 2021). The main reason for this is that the Basin largely corresponded to the historical territory of Hungary, consequently making it unnecessary to use an alternative name for the same entity. However, as during the 1910s, fears of potential territorial losses increased in Hungary, the need was increasingly felt to stress not just the environmental but also the political and economic coherence of the Basin. Accordingly, following the territorial losses in 1920, the term ‘Carpathian Basin’ experienced its heyday as a surrogate for pre-WWI Hungary, which Hungarian geographers and others made all effort to justify on all fronts (*ibid*). In fact, territorial revisionism was an official policy in interwar Hungary and was partly implemented during WWII. However, following the war, the reinstallation of interwar Hungary’s territory, and the installment of a communist regime, the ‘Carpathian Basin’ became a taboo term until the early 1980s (*ibid*). Since then, however, the concept has gradually sneaked back into Hungarian academia, public awareness, and eventually politics (Scott & Hajdú, 2022).

Although all governments have – to varying degrees – embraced the issue of trans-border Hungarians<sup>1</sup> in the Carpathian Basin since 1990 (Waterbury, 2010), 2010 certainly constitutes a milestone in this respect. In fact, Hungary’s since-then-incumbent national-conservative government coalition has made this issue one of its key ideological cornerstones (Lesińska & Héjj, 2021). At the same time, it is important to note that – unlike in the interwar period – the recent revival of Carpathian Basin-related discourses and policies are envisaged in line with respecting current borders and cooperation within EU frameworks (Bán, 2015). Still, it is on some level understandable that the concomitant narrative of ‘reunifying the nation’ (*ibid*, Pogonyi, 2017) raises eyebrows, especially among Hungary’s neighbors. To what extent the expression may be divisive in Hungary itself is the subject of investigation in the remainder of this paper.

<sup>1</sup> The literature uses various terms for Hungarians living beyond Hungary’s borders. Here, we refer to them as ‘trans-border Hungarians’ because this is one of the simplest and most used phrasings. Later in the paper, we also apply different terms like ‘Hungarians beyond the borders,’ ‘ethnic Hungarians,’ and ‘(ethnic) minority Hungarians.’



## 4 The parliamentary public sphere

Parliaments are among the most important institutions of representative democracy, being political assemblies that engage in free and open political debate on legislation. The key question in examining political representation is the extent to which the parliamentary agendas reflect the concerns of the public. Moreover, since communication is the core of parliamentary decision-making, it is also the core of representative democracy itself. At the same time, the parliamentary discourse is not a sterile reasoning process since representatives are aware that they cannot realistically convince their political opponents. Instead, politicians tend to emphasize symbolic issues and reinforce values that constitute political ideology. This is why the Parliament is also an excellent research field for us.

Discourses in parliaments not only reflect social configurations but also contribute to the discursive formation of them – in our case, through the use of geographical concepts with certain historical/cultural connotations.

Since the focus of our analysis is the language usage in parliamentary speeches, it is worthwhile discussing the discursive context briefly. Regarding the focus of the parliament, we distinguish between two operational categories, debate parliaments and working parliaments (Gallagher et al., 2011). The Hungarian Parliament – like most European parliaments – can be considered a working parliament, with less confrontational interactions, where the focus is on legislative procedures and committees. However, beyond legislation, the Hungarian parliament is the arena of governmental responsibility and accountability, where MPs confront cabinet members with current political issues and social problems, too (as happens in parliamentary government systems).

In parliaments, the power of speaking is the power of acting (Ilie, 2017): what can be done depends to a large extent on what can be said. This is also an important thought in relation to the changing legitimacy/role of the term ‘Carpathian Basin,’ which is detailed above. Speeches contain both theatrical and agonistic elements, i.e., they address both formal and competitive goals (Ilie, 2003). Since the Parliament in the Hungarian public law structure undertakes both legislative and control functions, the discourses can be distinguished on this basis, which is roughly in line with what Ilie states from a linguistic viewpoint: the genre of parliamentary discourse displays several subgenres that are subordinated to these two specific parliamentary goals. Keynote speeches, speeches, and pre-agenda speeches generally have a legislative function and, thus, a representative function that reflects the main messages of the parties. Speeches with a control function typically include prompt questions, prompt replies, two-minute speeches, etc., which by definition are non-representative, non-ceremonial speeches.

From a text analysis perspective, a further important genre distinction is whether the text is pre-written or spontaneous. This distinction is broadly consistent with the functional distinction above. Additionally, the debates are audience-centered as they occur in front of a real audience of other MPs and a virtual audience of voters and the media.

Our goal is also to examine the differences in the language of the parties in relation to the concept of the Carpathian Basin. General differences in the language of Hungarian parties have been investigated qualitatively by Szücs (2012). According to the latter, the Hungarian party Fidesz has created a ‘new conservative language’ as a successful alternative to the ‘sociologizing language’ of the 1990s. Szücs identifies distinctive features in

comparing the latter two languages, such as moralistic versus professional, referring to the natural order versus socioeconomic factors, rhetorical versus denying rhetoric, etc. These are the features we can expect to see in our analysis when comparing political blocs, even if, in contrast to our approach, Szücs observed the representative public sphere, not the Parliament.

## 5 Political landscape

Characterizing the time interval under review according to the parties in government, three periods can be distinguished: 1998-2002: Competition between center-right parties led to FIDESZ's leadership by center-right government; 2002-2010: The socialists (MSZP) and liberals (SZDSZ) twice became the governing force. During this period, FIDESZ became a mobilizing populist party, leading the FIDESZ-KDNP alliance to a landslide victory in the 2010 elections, resulting in a two-thirds majority in parliament. After 2010, Fidesz re-structured the whole Hungarian political scene: with a supermajority in the parliament, it adopted a new Fundamental Law and radically reorganized the whole electoral process, which resulted in not only successive electoral victories for the governing party but also a two-thirds supermajority in each parliamentary cycle.

For our analysis, we divided the period under consideration into two parts, with the 2010 elections as the cut-off point. As mentioned above, 2010 is a turning point for our topic. The right-wing Fidesz's return to power in 2010 gave them (and their election partner, KDNP) a constitutional majority in Parliament. Viktor Orbán, as new Prime Minister and former opposition leader, called for a 'revolution in the polling booths.' The party started implementing fundamental changes. The Constitution was replaced with a Fundamental Law, which states the emergence of a new order, the 'System of National Cooperation' (Nemzeti Együttműködés Rendszere). The regime's most important messages were based on the idea of ethnicity-based 'national (re)unification' (Pogonyi, 2017).

In the meantime, the role of the Parliament was also reduced. For instance, from 1999 onwards (during the first Fidesz government), the weekly plenary session was reduced to meetings every three weeks. After a change of government, between 2002-2010, bi-weekly sessions were the norm. Since 2010, when Fidesz returned to power, the former three-week system has been reinstalled. Moreover, the functioning of the parliament has changed radically: the lawmaking process has become less transparent with the reorganization of committees, and the opposition's rights have been severely reduced (Szente, 2020).

From the point of view of our topic, it is worth briefly reviewing the symbolic political points that have defined the relationship with Hungarians beyond the borders during the period under discussion. The first important event was the referendum on dual citizenship in 2004, the stake of which was whether Hungarians living beyond the border could be granted Hungarian citizenship. The real public debate around transborder Hungarians was invigorated around 2001 when the then-governing Fidesz-led coalition introduced and passed a law about the preferential treatment of Hungarian individuals beyond the borders (including financial aid, access to schools, cultural facilities, and work permits, among other elements). Fidesz openly refused to provide citizenship for trans-border Hungarians and offered this special status instead. However, this policy was not really



welcomed on the right: In 2003, a right-wing organization called Magyarok Világszövetsége (World Federation of Hungarians) initiated a referendum about granting a preferential process for obtaining citizenship to Hungarians beyond the borders. Though many found the idea risky, Fidesz (already in opposition) quickly joined the referendum campaign, claiming that this could not be a political issue and that refusing citizenship could not be an answer. The governing left-liberal coalition was strongly divided by the issue, but many of their leading politicians campaigned against the proposal, stating that dual citizenship was not a solution for these people (Bárdi, 2013). The harsh campaign and the failed referendum (neither option reached the required 25% threshold) stabilized the issue of Hungarians beyond the border as a kind of political cleavage and a political polarization factor. The next turning point was in 2010: among their first legislative acts after their supermajority win, Fidesz elevated the day of the Trianon Peace Treaty (1920, when Hungary suffered severe territorial losses) to a national Memorial Day, and a completely new national policy was launched. Since 2011, laws have been enacted to provide kin-state citizenship, voting rights, and financial and political support to ethnic Hungarians abroad, especially in neighboring countries. The issue of Hungarian minorities beyond the borders has divided the political parties in Hungary: while the left-wing and liberal parties promoted constitutional patriotism, the right-wing parties supported the idea of the 'virtual unification of the nation.' Achieving a constitutional majority in the Hungarian parliament opened the door for Fidesz to reshape the citizenship issue and provide extensive financial aid and other state-funded policies (schooling, social services, infrastructure development, etc.) to the communities of Hungarians beyond the border. As Kiss summarizes, the landslide changes in Hungarian politics fundamentally transformed so-called national policy, both politically and financially (Kiss, 2018, pp. 57-63).

Finally, in the context of the Carpathian Basin as a historical self-definition, it is also worth mentioning that after 2010, Fidesz became very active in instrumentalizing collective identity-making. By excluding opposition figures and their historical narratives from the public sphere, Fidesz presented itself as the sole defender of Hungarian national sovereignty. In this effort, the 'us/them' contrast has often been used: specific collective memories, including the Trianon Treaty, have been reinterpreted and used to assert a sense of national belonging and resentment against other nations (Benazzo, 2017).

Since our goal was primarily to detect political differences in the parliamentary discourse, we had to categorize parties according to their political stance. Members of the following parties were elected in the period under review: Alliance of Free Democrats (SZDSZ), Christian Democratic People's Party (KDNP), Democratic Coalition (DK), Dialogue for Hungary (PM), Fidesz – Hungarian Civic Union (Fidesz), Hungarian Democratic Forum (MDF), Hungarian Socialist Party (MSZP), Independent Smallholders, Agrarian Workers and Civic Party (FKgP), Jobbik Movement for a Better Hungary (Jobbik), Politics Can Be Different (LMP), Together – Party for a New Era (Együtt). The ideological classification of several of these parties is problematic. Fidesz has not been ideologically coherent in the period under study at all: until 2002, it pursued a basically conservative-liberal economic policy, then after 2002, in opposition, it took a left turn regarding economic policy, while in foreign policy it followed an Atlanticist and Westernist line, and finally, since 2010 it can be defined as a right-wing populist and national-conservative party. Furthermore, LMP claims to be green, but it is very difficult to locate them on the left-right scale.

In the end, we decided on three categories: Fidesz, other right-wing nationalist parties (Jobbik, MDF, FKGP, KDNP, MIÉP), and left-wing liberal parties (SZDSZ, MSZP, DK, LMP, PM, Együtt). Fidesz is listed separately as a benchmark. The latter two political blocs, especially the left-liberal one, are ideologically very heterogeneous: we do not define a political position associated with a clear-cut ideological stance but as a point of reference to Fidesz as the hegemonic party. This classification is based partly on the Parties and Elections Database (Nordsieck, 2022) and partly on the literature on the Hungarian party system and political cleavages (Horváth & Soós, 2015). According to the latter, instead of the sometimes confusing or overlapping ideological categorizations of the parties, the relationship between the parties (their cooperative or hostile attitude towards each other; in other words, the possibility of their coalition-making) may be a good means of categorization, as it indicates an important political cleavage since the mid-2000s when supporting the major parties and their leaders (like Viktor Orbán and Fidesz, or Ferenc Gyurcsány and the Socialists) also became a kind of political cleavage, beyond the classic ideological orientations (Horváth & Soós, 2015, p. 277). Furthermore, electoral cooperation during the period covered by the research also strengthened our bloc categories: far-right parties never formed an electoral coalition with the left, liberal, and green parties until 2022, and no party from the latter groups offered/accepted genuine cooperation from Fidesz, but always engaged in stronger or weaker coordination among themselves.

In the parliamentary discourses related to the Carpathian Basin, we expect to see an increase in the legitimacy awarded to and different meanings of the concept. Since the examined period parallels the birth of Fidesz's new, distinctive national policy and memory policy, it was assumed that this would also affect the discourses. Finally, in relation to the 'sociologizing language' of the left-liberal side, instead of emotional identification and the use of national symbols, a kind of bureaucratic narrative can be expected in the speeches of this bloc as well.

## 6 Data and methods

### 6.1 Data

Our corpus contains speeches made in the Hungarian Parliament from parliamentary terms 3-7 between June 25, 1998, and November 23, 2020. The start date marks the formation of the new Parliament, and the end date marks the end of our data collection. The openly accessible corpus was collected within the PARLDATA project of the non-profit organization K-Monitor by its volunteer developers and the consulting firm PrecognoX. We restricted the analysis to the relevant speech types by eliminating non-substantive, technical ones, e.g., those related to the agenda. Only speeches of party-affiliated MPs were analyzed, so independent MPs and representatives of national and ethnic minorities were excluded from the dataset, resulting in a corpus of 168,506 speeches.

The relevant sub-corpus was identified by keyword filtering, keeping only speeches containing the term 'Kárpát-medenc\*' (the Hungarian equivalent of Carpathian Basin; \* replaces any string here; this is necessary due to the agglutinative nature of Hungarian). Our final corpus consists of 1525 speeches, and our unit of analysis is full speeches.

## 6.2 Methods

The purpose of topic modeling is to uncover hidden topics in a corpus. The model assumes the existence of a finite set of topics, where a topic is statistically defined as a multinomial distribution over the specified parliamentary terms. The model allows texts to relate to more than one topic. Structural topic models (STM, Roberts et al., 2019), a subtype of topic models, allow the researcher to estimate the relationship of topics to document metadata. This metadata can influence both the content and prevalence of topics, where content refers to the words used within a topic. The metadata option of STM fits well with our research question because it allows us to examine how political blocs speak differently about a given topic. Further, the vocabulary differences between political blocs may indicate a divergence in the framing of the topics concerning how political blocs interpret and communicate political reality. In fact, STM was developed by political scientists and adapted for purposes similar to ours, such as examining vocabulary differences by party (Roberts et al., 2019).

Another important feature of the model is that, in contrast to classical topic models (Blei et al., 2003), it allows for correlation between topics. It does not make the rather unrealistic assumption that topics that appear in individual speeches are independent. To implement the models, we used the STM R package (Roberts et al., 2019).

Like many other text-mining approaches, STM relies on the word-based representations of texts (e.g., Aggarwal & Zhai, 2012), assuming that texts are ‘bags of words’ while ignoring word order, syntactic relations, etc. During the preprocessing, we unified the character encoding of the database and removed unnecessary tokens that had been added to the text during the web harvesting process. We employed lemmatization to standardize different forms of the same word. For this purpose, we used the e-magyar (emtsv package (Váradi et al., 2018; Indig et al., 2019)) and deleted very common words (‘stop words’ like articles and conjunctions). We treated the most relevant two- and three-word collocations (‘significant bigrams/trigrams’) as single terms, such as ‘*határon túli magyarok*’ (‘Hungarians beyond the border’) or ‘*előző ciklus*’ (‘previous term’). We recognized proper nouns (or ‘named entities’ in technical terms), such as names of politicians, parties, etc., and treated them as single terms.

The number of topics is an input parameter of the STM model and the metadata variables together with their type (‘content variable’ or ‘prevalence variable,’ depending on whether they affect the content or prevalence of the topics, with the restriction that only one content variable can exist). When trying to optimize the choice of these input parameters, our decision was based on the interpretation of the topics of the models obtained from different inputs.

As mentioned above, we used two meta-variables, date, and political position. The date (as previously explained) was binary: before/after the 2010 elections, while the political position was a three-category variable: Fidesz/right-wing nationalist bloc/left-liberal bloc.

When deciding on the type of meta-variables, we fitted two different seven-topic models. A topic size of seven was considered realistic based on our previous modeling experience with the corpus. Both models had political position as a content variable, but in one, only the date; in the other, the political position was also a prevalence variable. From the point of view of interpretability (displaying words associated with topics and meta-variables, using the `sageLabels` and `labelTopics` function of the STM package), the latter model seemed to be better, so this meta-variable representation was chosen.

To determine the optimal number of topics, we combined quantitative and qualitative approaches to get topics that were not too broad but still not over-clustered. First, we fitted the model according to varying numbers of topics (between 5 and 20), and assessed the models using the searchK function, which performs several tests. This data-driven approach helped us to identify a narrower range of topics, 7-11. Inspecting further now only within this narrower class of models, according to their most relevant terms, topic sizes eight and nine seemed to be the most coherent and, simultaneously, the most parsimonious.

We ranked these two models qualitatively by reading the ten speeches most strongly associated with each topic in six groups (two periods and three political blocs) using the findThoughts and plotQuote functions. Finally, a topic size of eight proved to be the best for meaningful interpretation.

### 6.3 Limitations

Our analysis, like any other, has its limitations. It examines political communication only in Parliament, a medium characterized by a specific genre of discourse. Additionally, our unit of analysis was whole speeches, although we could have chosen a smaller context for the occurrence of the term 'Carpathian Basin.' However, we felt that the whole speech was representative of the discourse in which the term occurs. Furthermore, although the parliamentary speeches could be considered elements of broader debates with dynamic and interactive patterns, this research analyzed the speeches as individual manifestations.

Though the issue of the Carpathian Basin, both as a territory and a metaphor for Hungarians living there, was an important issue in the early years of the 1990s in Hungarian politics, as the scope of the research covers the 1998-2020 timeframe, the paper does not deal with earlier political developments. We divided the examined period into two intervals. A finer division could have been employed, but the 2010 split point reflected the theoretical aspects. Finally, since there are no canonical procedures for determining the optimal topic model, quantitative and qualitative approaches were used to select the model parameters.

## 7 Results

### 7.1 Descriptive results

The number of speeches using the term 'Carpathian Basin' increased significantly in the period under review (Table 1; note, however, that the second interval is shorter). If we look at the occurrences of the term rather than the speeches containing the term, the change is negligible (it decreases from 86,743 to 81,763). This may indicate that the number of contexts/areas in which the term is used is what is actually increasing.

The increase in the number of speeches was not uniform across political blocs. The number of speeches using the term tripled for Fidesz and the right/conservative bloc, while the number of speeches on the left-liberal side was unchanged (Table 1). In general,

it is also true that most of the speeches were made by Fidesz representatives, and the left-liberal side made fewest. The same pattern holds for the proportion of speeches containing the phrase as a percentage of all speeches: left-liberal speeches contain a much smaller proportion of the keyword.

**Table 1** Number and proportion (%) of speeches containing the term ‘Carpathian Basin’

Speeches containing the term				
Date	Speaker’s political bloc			
	right-wing nationalist	Fidesz	Left-liberal	Total
1998–2010	153	161	161	475
	32.21%	33.89%	33.89%	100%
	27.82%	24.73%	46.69%	31%
2010–2020	397	490	163	1050
	37.81%	46.67%	15.52%	100%
	72.18%	75.27%	50.31%	69%
Total	550	651	324	1525
	36.07%	42.69%	21.25%	100%
	100%	100%	100%	
All speeches				
1998–2020	41,538	57,390	69,578	168,506
	24.7%	34.1%	41.3%	100%
Proportion of Carpathian Basin-related speeches within all speeches				
1998–2020	1.3%	1.1%	0.5%	168,506

‘All speeches’ refers to the corpus before keyword filtering, i.e., all other cleaning/filtering has been executed

Compared to the whole corpus, the Carpathian Basin sub-corpus is associated with a much higher frequency of pre-written keynote speeches, speeches, and pre-agenda speeches that have a legislative function in representing the main messages of the parties and are written in advance (see the Appendix, Table A1). At the same time, the sub-corpus has a much smaller proportion of non-representative, non-ceremonial, spontaneous speeches with a control function (prompt questions, prompt replies, two-minute speeches, etc.).

## 7.2 Modeling results

The final eight-topic model is presented in Table 2. The table presents the topic labels obtained after qualitative and quantitative interpretation, the prevalence of the topics, and their most relevant terms. The robustness of the eight-topic model (and the validity of the method in general) is demonstrated by the fact that the interpretation of the seven-topic and nine-topic models (which were fitted independently to the eight-topic one) shows that the topics can be matched (e.g. the topic of collective identity-making is present in each model), and a topic from the smaller model was split to obtain the larger model (e.g. the topic of climate and energy policy of the nine-topic model was derived from the topic of economy in the eight-topic model).

Prevalence (here and hereafter) is defined as the average of the topic contributions. To make sense of ‘contributions,’ recall that topics are not clusters of speeches since a speech can belong to more than one topic. The value of a topic’s contribution to a speech tells us the extent to which the given topic contributed to the ‘generation’ of the speech.

The most relevant terms are defined by FREX (Roberts et al., 2019). That is, they are not the most frequent terms within the topic but the most distinctive ones: FREX weights words by their overall frequency and how exclusive they are to the topic. Most of the eight topics are semantically well defined; Topics 6 and 8 are rather exceptions, with less homogeneous themes and typically lower topic contributions (if the highest values of these contributions are less than 80-90 percent, this indicates that no speech clearly belongs to the topic; the topic is typically mixed with other topics).

Correlation between the topics was not found to be considerable (absolute value below 0.25), so there is no relation between their occurrence in speeches.

We can describe the topics in more detail using a qualitative approach by reading the most representative speeches. We have focused on the ten most typical posts (with a topic contribution of at least 80%) per topic, per time period, and per ideological bloc. The most relevant examples of the use of the term ‘Carpathian Basin’ from these speeches, with short quotations, are included in Table A2 in the Appendix, from which the following quotations are taken. Table A2 also includes the original Hungarian text.

Topic 1 mainly includes speeches on agriculture and related sectors (water management, environment- and climate policy, sustainable energy), and the Carpathian Basin is presented as a transboundary ecosystem, e.g., in terms of GMO exemptions or pesticide use (‘in fact, water from all over the Carpathian Basin flows here.’). Elsewhere, the use of the space category seems less justifiable – for example, the climate/agro-economic potential of the Carpathian Basin is identified with Hungary’s climate/agro-economic potential, etc. For example, ‘Forest management in Hungary has been based on ten-year plans for 130 years. The result of this 130 years of planned work is, in my opinion, the survival of the forests of the Carpathian Basin.’ On the right-wing nationalist side, the debate about policy issues is not free from assessment of the geopolitical situation (‘[...] we can also say that the quality of life in this endangered place is primarily up to us; to us and to some extent, of course, to the neighboring countries, because in the current political situation, the whole Carpathian Basin is not under our jurisdiction.’)



**Table 2** Topics in the final model, with their label, prevalence, and most important terms

ID	Label	Prevalence (1998–2010, 2010–2020)	Most relevant terms
T1	Agriculture	14% (15%, 14%)	víz (water), termőföld (farmland), növény (plan), GMO, árvíz (floods), hektár (hectare), mezőgazdaság (agriculture)
T2	Culture	8% (5%, 10%)	hungarikum (unique Hungarian product), palinka (traditional fruit spirit), emlékév (commemorative year), termék (product), értéktár (heritage), világörökségi (world heritage), érték (value)
T3	Public administration, political rights	16% (16%, 16%)	ön (you), szavaz (vote), Jobbik (Jobbik, right-wing political party), státustörvény (status law), választás (election), Fidesz (Fidesz, currently governing party), párt (party)
T4	Collective identity-making	17% (18%, 16%)	nép (people), király (king), korona (crown), szabadság (freedom), történelem (history), Szent Korona (Holy Crown), hős (hero)
T5	Economy	10% (11%, 9%)	költségvetés (budget), százalékos (in percentage), demográfiai (demographic), stratégia (strategy), infláció (inflation), éghajlatváltozás (climate change), Duna (Danube)
T6	Human resources I: symbolic elements and institutions	7% (6%, 8%)	nemzetiségi (person belonging to a nationality), nemzetiség (nationality), műemlék (monument), egyház (church), roma (roma), kulturális örökség (cultural heritage), nemzetiségi önkormányzat (national minority government)
T7	Elements of national policy	19% (26%, 16%)	autonómia (autonomy), nemzetpolitika (national policy), határ túli Magyar (Hungarians beyond the borders of the country), Románia (Romania), szülőföld (homeland), Székely (Székely, family name or an ethnicity in Romania), csatlakozás (accession)
T8	Human resources II: social benefits, state aid, and funds	8% (3%, 11%)	Erzsébet (Elisabeth; The 'Erzsébet Camps' program offers holidays for children, including those from beyond the border), tábor (camp), alapítvány (foundation), millió forint (million Hungarian Forints), Bethlen Gábor Alap (Gábor Bethlen Fund, which supports cross-border Hungarians), BGA (abbreviation of Gábor Bethlen Fund), milliárd forint (billion Hungarian Forints)

Topic 2 includes cultural topics focusing on the whole Carpathian Basin, such as the ‘hungarikum’ label (which refers to high-quality local products), language law, or public television broadcasting beyond the borders. Here, the geographical label is usually used in the sense of ‘the Hungarians of the Carpathian Basin,’ but even here it is sometimes less justifiable – e.g., family-centricity is referred to as a traditional value in the Carpathian Basin (‘In the Carpathian Basin, people have been brought up within the family for centuries in a way that [...]’), or in speech intended to mobilize the Carpathian Basin (‘to listen, to mobilize the country, the Carpathian Basin in this matter’). Regarding this topic, the right-wing nationalist bloc also uses wording that suggests criticism of the geopolitical situation: ‘[it] is never a foreign country for us, the occupied Hungarian territory of the Carpathian Basin, any Hungarian territory, but formally, legally a foreign country.’ There are cases when the term is used in a neutral way, referring to an actual geographical region; these only occur in left-liberal speeches.

Topic 3 includes public administration and regulatory issues affecting Hungarians beyond the borders, such as voting rights, transport development, health care, and housing subsidies. This topic is rather interactive and contains controversial issues, as also indicated by the use of ‘you’ among the relevant terms and ‘not angry’ (a lemmatized form of the phrase ‘don’t be angry, but...’) among the most frequent terms. In often heated debates, the concept is also used in extreme contexts, e.g., ‘[...] we are dealing with the number one lobbyist of open society, who is completely indifferent to whether or not there will be Hungarians living here in the Carpathian Basin in a hundred years’ time.’

Topic 4 contains speeches related to collective identity-making, e.g., concerning the 1100th anniversary of the founding of the state. On the one hand, the Carpathian Basin as a whole is presented as the territory of the former Hungarian Kingdom, and on the other hand as the historical home of the Hungarians for thousands of years, where (according to the right-wing nationalist narrative) other peoples were either dispersed or merged into other nations. Here, too, the two geographical entities, Hungary and the Carpathian Basin, seem to overlap in some places (for example, ‘[...] to address the National Assembly, the people of the country, the people of the Carpathian Basin, the Hungarian nation all over the globe’). On the left-liberal side, the idea of the Carpathian Basin’s superiority is criticized, with the suggestion that we do not live alone in the Carpathian Basin (‘We have not lived our difficult history, this millennium alone, and we do not live alone in the Carpathian Basin today.’). In this historical context, the expression ‘the peoples of the Carpathian Basin’ is common from this political side, while on the other two sides, the Carpathian Basin is often referred to as a ‘historic homeland.’

The relevant terms in Topic 5 are all associated with the economic field, and the Carpathian Basin context is only one aspect of more general economic topics (budget, organizations’ annual reports, etc.) – e.g., mentioned as part of plans for an agricultural development or higher education development program for the whole Carpathian Basin.

Topics 6 and 8 deal with policy areas other than the economy, agriculture, and culture. Topic 6 deals with symbolic elements and institutions (church policy, monument protection, sports policy, etc.), and Topic 8 deals with social benefits, state aid, and funds. Both topics are rather heterogeneous, and typically, as in Topic 5, the Carpathian Basin arises in relation to these more general topics, alongside several other aspects (Topic 6: ‘the wooden churches of the north-eastern Carpathian Basin’ or the ‘winemaking



community in the Carpathian Basin,' Topic 8: '[...] which would make it possible to increase the knowledge about Hungarians living beyond the borders in the Carpathian Basin through class excursions [...]').

Finally, Topic 7 is the most interesting from a social research point of view because it also presents principled positions on national and neighborhood policy, typically along the lines of issues related to the political, cultural, and religious organizations of Hungarians living beyond the borders. It emphasizes symbolic community with Hungarians beyond the borders, which, according to left-liberal MPs, can be achieved by joining the EU ('It is fundamental for the Hungarians living in the Carpathian Basin that the process of reuniting the divided Hungarians through European integration without changing borders will continue and, we hope, be completed.'), while right-wing nationalist parties often associate this unification with national autonomy ('Hungarian autonomy aspirations in the Carpathian Basin,' and 'autonomy policy in the Carpathian Basin.'). The desire to create institutions for the symbolic reunification of the nation is often explicit in Topic 7 – for example, from a Fidesz MP: 'in the objectives of the government of the civic coalition, more systematic, closer, so to speak institutionalized, relations between the Carpathian Basin and the Hungarians of the West and the motherland were also a priority.'

In a speech pertaining to this topic, a right-wing nationalist MEP<sup>2</sup> explicitly referred to the national-political connotations of the term 'Carpathian Basin' and its advantages over other alternatives: 'So when we talk about Hungarians beyond the borders, we should always put this adjective in brackets; let's talk about Hungarians, let's talk about Hungarians living in the Carpathian Basin and in the world, let's talk about our annexed or separated brothers and sisters, but I think that we cannot go in the direction of what, among others, the policy of abandoning national interests has embodied [...].'

As regards the dynamics of the topics, the biggest change is the decrease in the prevalence of symbolic topics (see Table 2, third column: national policy from 26 to 16 percent, but also memory policy from 18 to 16 percent), and the parallel increase in the prevalence of public policy topics that replace symbolic gestures with concrete action in relation to their subject, e.g. financial support for Hungarian organizations beyond the borders (social benefits, state aids and funds from 3 to 11 percent, symbolic elements and institutions from 6 to 8 percent, cultural policy from 5 to 10 percent). If we examine these dynamics by political bloc (see Table A3 in the Appendix), we see that the decrease in the prevalence of symbolic topics mainly occurs with the left-liberal bloc, while the increase in concrete actions is mainly induced by Fidesz (in government).

Looking at the topics' relation to speech genres (see Table A4 in the Appendix), we see a big difference mainly in the pre-agenda speeches (which are, as mentioned, over-represented in the corpus under study). This is a genre that has a strong political and social significance; it thematizes the parliamentary session and the political public sphere. In this genre, the two symbolic topics are over-represented: memory policy and national policy (31 and 28 percent, respectively, compared to only 17 and 19 percent in the whole Carpathian Basin-corpus). In contrast, Topic 3 contains more speeches with a control

<sup>2</sup> Dr. Tamás Gaudi-Nagy, from the far-right party Jobbik, 2013

function, which are rather non-representative, spontaneous speeches. This is consistent with the interpretation of the topic above, which presents it as a more interactive, conflictual topic.

Finally, let us see how different political blocs talk differently about the same topic. The plots below show which words within a topic are more associated with one political bloc versus another. Due to the language of the corpus, the terms in the figures are in Hungarian. The horizontal axis quantifies the degree of belonging to the blocks; the colors orange, dark brown, and blue represent the Fidesz, the right-wing nationalist bloc, and the left-liberal bloc, respectively. In Figure 1, on the left side of the plot, the words on the left are those that Fidesz uses more frequently within Topic 1, while on the right side of the plot are words that the left-liberal block uses more frequently. The size of words is proportional to their frequency in the corpus. The figures were originally plotted using the `plot.STM`, `type = 'perspectives'` function, then recreated in Tableau. Below is a selection of just a few plots that highlight important results, but the online supplement includes an interactive visualization that allows users to explore all the plots by selecting various parameters.

In several topics (e.g., T1 and T4, see Figures 1 and 2), the general difference is that while Fidesz and right-wing nationalist MPs speak in terms of representing their own community (the pronoun 'we' is over-represented) – i.e., forming an 'us'/'them' opposition – the left-liberal side is more formal, less personal and refers to institutions ('EU', 'government', 'Hungary'). A similar contrast is reflected in the collective identity-making topic (T4, Figure 2), with Fidesz and the right-wing nationalists naming their own community on an ethnocultural basis ('nation', 'Hungarian') and the left naming its own community on a civic basis ('Hungary'). In T1 on agriculture, the two large blocs are also markedly differentiated by the fact that the two right-wing blocs also include a territorial aspect (the word 'territory' is over-represented).

The difference between discourse referring to one's own community and to organizations can also be seen as the difference between the homogeneity and heterogeneity of the perspectives that are represented, most clearly in the case of Topic 7 (national policy, see Figure 3). Here, the narrative of Fidesz and the right-wing nationalists is dominated by their own community defined on ethnocultural grounds ('Hungarian', 'Hungarians',<sup>3</sup> 'community', 'national', 'connection'), while on the left-liberal side, the actors and perspectives in the narrative are more heterogeneous ('European', 'European Union', 'Hungarians beyond the borders', 'Hungary'). It is revealing that the Fidesz/right-wing nationalist half of both plots are dominated by a single term, 'Hungarian.'

A further dimension of the right-left divide appears in Topic 5 on the economy (see Figure 4): the two right-wing blocs argue on a value and emotion basis, while the left argues on a rational basis, the former on a rhetorical, the latter on a bureaucratic/expert basis (Fidesz vs. left-liberal: 'strategy', 'important', 'national', 'goal' versus 'budget', 'resource', 'government', right-wing nationalist versus left-liberal: 'Hungarian', 'family' versus 'government', 'budget'). The same distinction can also be observed in Topic 6 (Figure 5: Fidesz

<sup>3</sup> The Hungarian original ('magyarság') is a difficult-to-translate term that refers to the ethno-culturally homogeneous Hungarian nation as a whole living anywhere in the world.

versus left-liberal: ‘culture,’ ‘national,’ ‘support,’ ‘important,’ ‘Hungarian’ versus ‘economic,’ ‘interest,’ ‘law,’ right-wing nationalist versus left-liberal: ‘community,’ ‘culture,’ ‘nationality’ versus ‘interest,’ ‘strategy,’ ‘law,’ ‘government’).

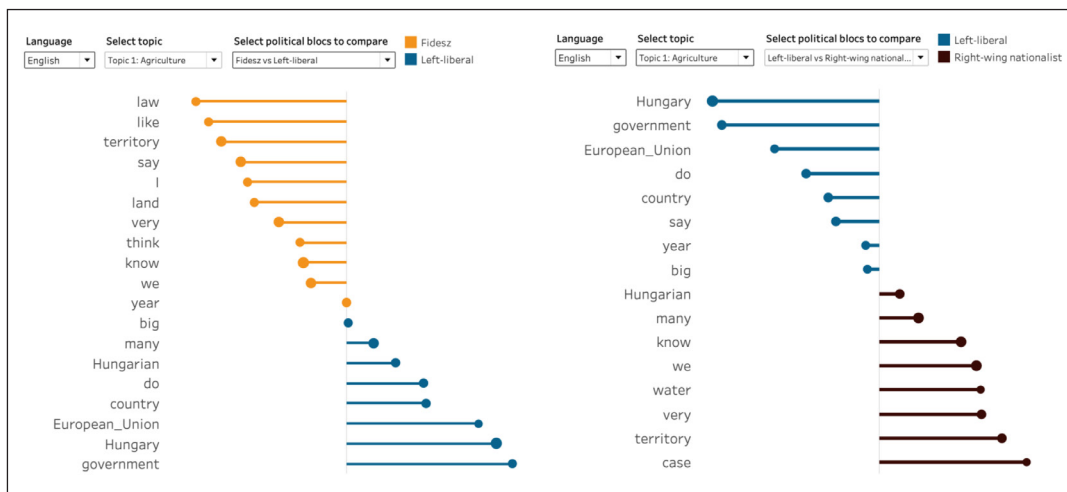


Figure 1 Vocabulary differences according to political blocs for Topic 1 (Agriculture).

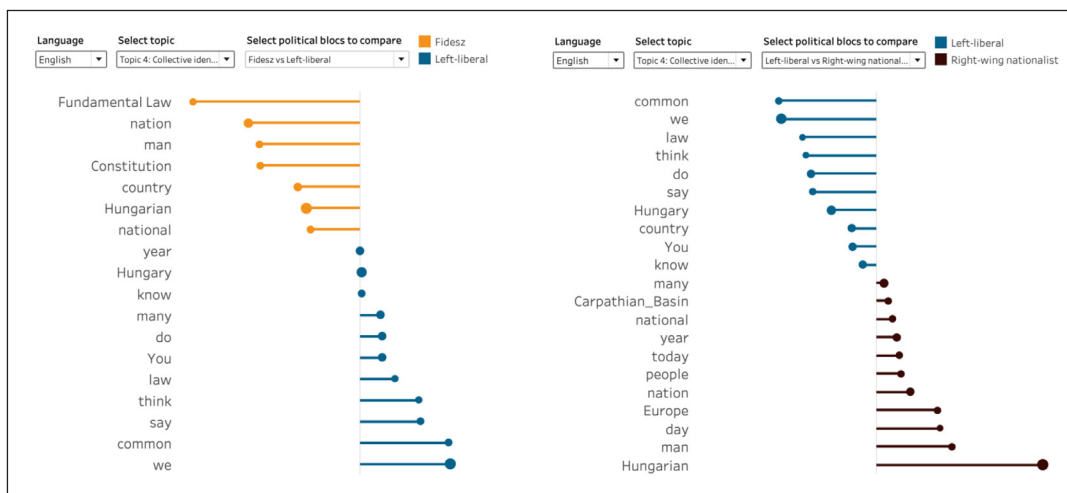


Figure 2 Vocabulary differences according to political blocs for Topic 4 (Collective identity making).

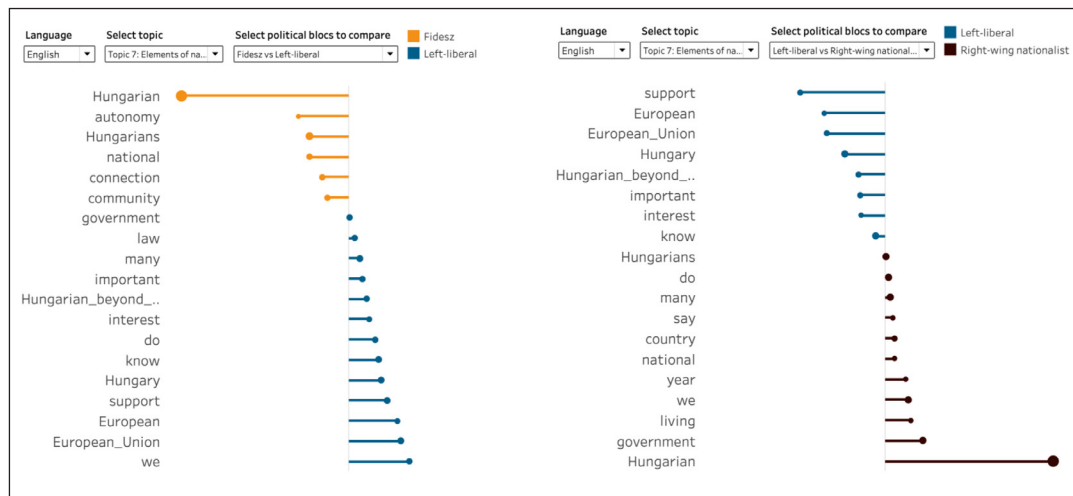


Figure 3 Vocabulary differences according to political blocs for Topic 7 (Elements of national policy)

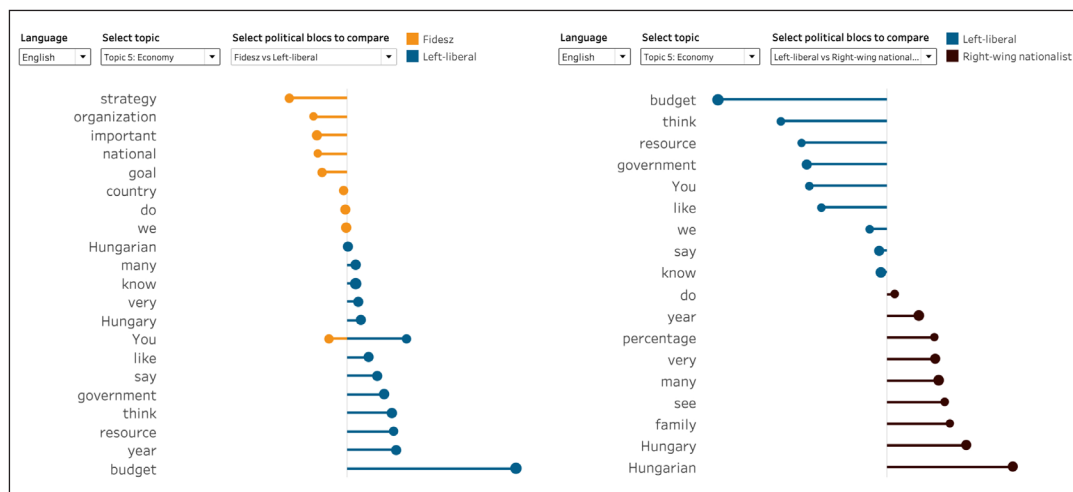


Figure 4 Vocabulary differences according to political blocs for Topic 5 (Economy)

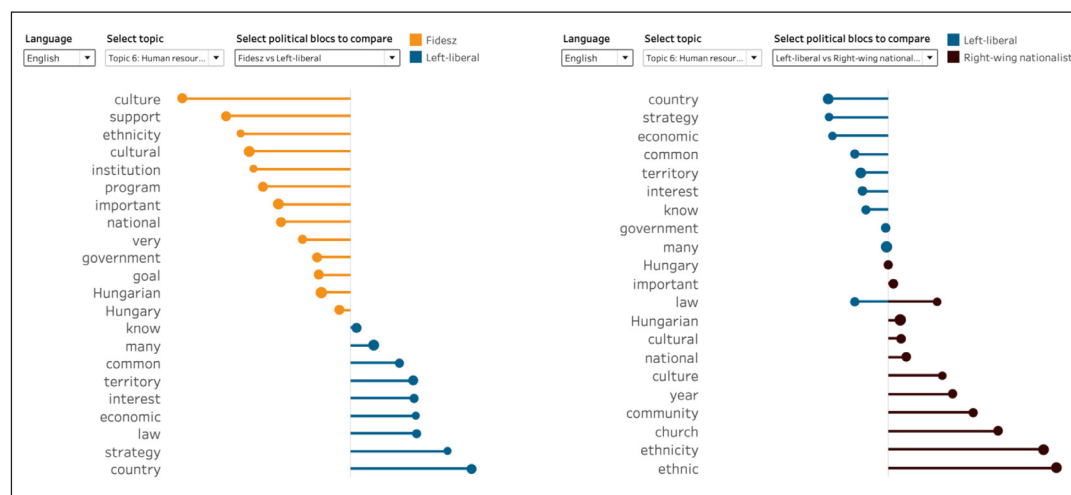


Figure 5 Vocabulary differences according to political blocs for Topic 6 (Human resources I: symbolic elements and institutions)

## 8 Summary and discussion

In our paper, we have examined how the Carpathian Basin, one of the most important geographical concepts related to national identity, was referred to in the past two decades in the Hungarian Parliament. We sought to answer the question of what latent themes can be distinguished in the discourses and what linguistic representation of political processes be captured.

Our analysis, like any other, has its own limitations, which are listed in the chapter on limitations. Here, we would like to focus on just one of them: our study examines political communication only in a specific political institution, the Hungarian Parliament. There is no available research on how Hungarian parliamentary discourse relates to, influences, or shapes the broader political communication space; existing analyses focus mainly on campaign communication, media relations, and changes in the political agenda (Kiss, 2019), and more recently, especially on online political communication (Bene, 2020). In the relevant research, parliament is usually considered an actor (legislator), not a layer of political communication. Research focusing on the political discourse finds that there is a sovereignty/integration dichotomy in the Hungarian political discourse that may be related to the context of territoriality, which is the topic of this paper. The European integration-oriented political discourse dominant in the 1990s has been complemented by a national sovereignty-oriented discourse since the 2010s (Szűcs, 2015). In our study, which examines the rise and political characteristics of the Carpathian Basin as a geographical and political category used in parliament, we did not use this sovereigntist-integrationist dichotomy as an analytical framework, but it is important to highlight the temporal coincidence between the rise of the sovereigntist discourse and the dominance of the term 'Carpathian Basin.' In this sense, our analysis can also be interpreted as quantitative confirmation of previous qualitative research.

Turning to our results, the political relevance of our topic is shown by the fact that while the number of related speeches in the left-liberal bloc remained unchanged during the period, the number of speeches in the Fidesz and the right-wing nationalist blocs more than doubled. These results can be added to the observation that the left-liberal side lost the opportunity to shape identity and memory politics in Hungary (Barna & Knap, 2022). The marked rise of the Carpathian Basin as a discourse unit in the Right's speeches after 2010 testifies to the sharp change in national policy and can be seen as a kind of political symbol of the renewed relationship between the Hungarian state and the Hungarians living beyond the border (Kiss, 2018).

The sub-corpus of parliamentary discourse related to the Carpathian Basin represents different genres than the full corpus: the former includes mainly representative ceremonial speeches conveying the main messages of the parties and performing a legislative function. All this may indicate that the term 'Carpathian Basin' is perceived by MPs as having a strong political meaning and is consciously used to convey a political message. However, the term may not yet be part of everyday vocabulary because it is less frequently used in spontaneous speeches. The fact that this topic is hardly present in the control-function speeches may also indicate that although national policy and politics related to trans-border Hungarians (especially concerning financial support and the government's political activities beyond the border) have been hot topics in Hungarian politics since 2010, opposition parties did not problematize this in the parliament, thus the politicians of the governing party did not mention it either in their replies. These are the linguistic imprints of the way in which the issue of Hungarian minorities beyond the border has divided political parties.

According to the results of structural topic modeling, we identified eight topics in the corpus which are well-separated thematically (economy, agriculture, culture, collective identity-making, elements of national policy, public administration and political rights, social benefits/state aid/funds and symbolic elements/institutions). The symbolic function of the topics on national policy and collective identity-making is clearly shown by the fact that these speeches were over-represented among pre-agenda speeches (which genre, in general, has a representative and opinion-forming function).

As for the dynamics of the topics, the left-liberal bloc is identified with a decrease in the prevalence of symbolic issues (collective identity-making, national policy), while the right-wing nationalist bloc saw an increase in the prevalence of public policy issues, which, in fact, replace symbolic gestures with concrete actions, such as financial support for Hungarian organizations beyond the borders. The results reflect the fact that Fidesz's rise to power has fundamentally transformed national politics, not only politically but also financially.

In the eight topics, the term is used in different contexts, often with different meanings. These contexts can be divided into two larger categories. In the first category, the term is sometimes used to refer to a politically neutral geographical unit (e.g., a trans-boundary ecosystem related to the topic of agriculture). Another type of use in this category is where the geographical unit's relation to Hungarians is more precisely specified ('Hungarian-inhabited areas of the Carpathian Basin'), or it is explicitly designated as an ethnically heterogeneous area and other actors are also explicitly mentioned ('nations of the Carpathian Basin,' 'peoples of the Carpathian Basin'). Left-liberal speeches typically

fall into this category. In the second category, the two geographical entities, Hungary and the Carpathian Basin seem to overlap: the agro-economic potential of the Carpathian Basin is identified with Hungary's agro-economic potential, family-centricity is presented as a traditional value in the Carpathian Basin, etc. The Carpathian Basin is often identified with historical Hungary; e.g., in speeches on the topic of collective identity-making, the Carpathian Basin has been presented as the historical home of the Hungarians for over a thousand years. This second category of use is typical of Fidesz and the right-conservative bloc.

The politically non-neutral use of the term can also be approached in relation to the 'virtual nationalism' concept of Csergő and Goldgeier (2004). Virtual nationalism entails a national agenda that aims at the integration of political unity with the national community, yet it rejects territorial claims and aims to establish institutions that enable the sustenance of a national community separated by borders. Some of the topics (Topics 3, 5, 6, and 8) relate to various public policy strategies that aim to establish these institutions. The purpose of these institutionalization efforts is often explicitly mentioned in Topic 7, which deals with the principles of national policy. The explicit distinction between the demand for territorial unification and the non-territorial means of nation-building is most pronounced on the left-liberal side; see the speeches that emphasize a symbolic community that can be achieved by joining the EU. On the other political sides, this distinction is sometimes not explicit; the unification is often associated with ethnic autonomy, and sometimes, the wording of right-wing nationalist MPs even suggests a denial of the legitimacy of territorial division, which means they go beyond virtual nationalism.

Different political blocs frame the same topics in different ways. The different vocabularies suggest that Fidesz and right-wing nationalist MPs speak in terms of representing their own community, forming an 'us'/'them' opposition, and defining their own community on an ethno-cultural basis. In contrast, the left-liberal side speaks in a less personal voice and refers to institutions, defining the community on a civic basis. The differences between left and right in the discourse on the territoriality of Hungary have already been presented in previous analyses: the Hungarian parliamentary discourse on the Treaty of Trianon between 1990 and 2002 shows that while the right-wing parties treated the peace treaty and its territorial consequences as a national-historical issue (tragedy), the left-wing parties focused on the social effects and consequences of the peace treaty and the loss of territory (Romsics, 2006). In other words, the political meaning of Trianon was very different for different actors, as our research on the Carpathian Basin also confirms.

The difference can also be found in terms of the homogeneity and heterogeneity of the perspectives that are represented. A further dimension of the right-left divide appears in the values vs. interests, emotions vs. bureaucratic expertise, culture and community vs. strategy, and government distinctions. This division is a long-standing feature of Hungarian political discourse, where the parallel existence of ideas about the cultural nation and the political nation characterizes the political discourse and policymaking of national policy (Vida, 2002).

The latter results are in line with Szücs' (2012) qualitative study on other genres of political texts (representative public speeches), in which he identifies the Fidesz/left-liberal distinction along the moralistic-vs-professional-, and in terms of justification according to the natural-order-vs-socio-economic-factors dimension. We can also put our results in a more general, international context, such as corresponds with Mouffe's influential writing



(e.g., Mouffe, 2011) on the current state of democracy. Mouffe argues that socio-democratic/liberal thinking ignores fundamental, contradictory aspects of human nature, and instead of creating a public sphere that allows conflicting values and interests to clash, its proponents believe in a universal rational consensus. They ignore the fact that there is an important affective dimension in political acting and that people need to be able to identify with a collective identity that gives them an image of themselves to which they attach values. In contrast, populist politics is characterized by an emphasis on emotion rather than rationality, the expression of collective identity rather than the principle of competing individuals. These dichotomies recur in the results of our analysis.

In sum, the term 'Carpathian Basin' used to legitimize official claims to territories lost by Hungary in the interwar period and hence not used on official forums for decades after the Second World War and even after the regime change, is now widely used by the Hungarian political elite on all political sides. However, the linguistic features of the parliamentary discourse related to the concept reflect the fault lines of Hungarian national policy and Hungarian political thought in general. In the left-liberal bloc, the term tends to be used to refer to a politically neutral geographical entity or to an ethnically heterogeneous area. It is mainly signs of virtual nationalism that can be detected in the politically non-neutral use of the term, but there are speeches on the right-wing nationalist side that go beyond this and implicitly suggest questioning the legitimacy of territorial division. The right-wing narratives, although in different ways and to different degrees, share the feature of demarcating the Carpathian Basin as a single geographical entity. It is presented as a Hungarian space that manifests itself in symbolic memory and national policy but is also strengthened through institutional and economic ties.

## Acknowledgment

We would like to thank Jakab Buda and Árpád Knap for their invaluable input throughout the data cleaning process.

## Funding

Eszter Katona's work was supported by the ÚNKP-23-4 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development, and Innovation.



## References

- Aggarwal, C. C., & Zhai, C. (2012). An introduction to text mining. In *Mining Text Data* (pp. 1–10). Springer US.
- Balogh, P. (2021). The concept of the Carpathian Basin: its evolution, counternarratives, and geopolitical implications. *Journal of Historical Geography*, 71, 51–62. <https://doi.org/10.1016/j.jhg.2020.12.003>
- Balogh, P. (2022). Spatial identity politics and the right in Hungary. In Mörner, N. (Ed.), *The Many Faces of the Far Right in the Post-Communist Space* (pp. 100–105). Huddinge: CBEEs, Södertörn University.
- Barna, I., & Knap, Á. (2023). Analysis of the thematic structure and discursive framing in articles about Trianon and the Holocaust in the online Hungarian press using LDA topic modelling. *Nationalities Papers*, 51(3), 603–621. <https://doi.org/10.1017/nps.2021.67>
- Bassin, M. (2012). National Metanarratives after Communism: An Introduction. *Eurasian Geography and Economics*, 53(5), 553–556. <https://doi.org/10.2747/1539-7216.53.5.553>
- Bán, K. (2015, July 25). A nemzet újraegyesítése a cél. *Magyar Hírlap*. [https://www.magyarhirnap.hu/belfold/A\\_nemzet\\_ujraegyesitese\\_a\\_cel](https://www.magyarhirnap.hu/belfold/A_nemzet_ujraegyesitese_a_cel)
- Bárdi, N. (2013). Magyarország és a kisebbségi magyar közösségek 1989 után. *Metszetek*, 2(2–3), 40–79.
- Benazzo, S. (2017). Not all the past needs to be used: Features of Fidesz's politics of memory. *Journal of Nationalism, Memory & Language Politics*, 11(2), 198–221. <https://doi.org/10.1515/jnmlp-2017-0009>
- Bene, M. (2020). *Virális politika. Politikai kommunikáció a Facebookon*. L'Harmattan Kiadó.
- Blei D.M., Ng A.Y., & Jordan M.I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Csergő, Z., & Goldgeier, J. M. (2004). Nationalist strategies and European integration. *Perspectives on Politics*, 2(1), 21–37. <https://doi.org/10.1017/s153759270400060x>
- Dodds, K. (2019). *Geopolitics: A very short introduction (3rd ed.)*. Oxford University Press.
- Egry, G. (2020). The greatest catastrophe of (post-)colonial Central Europe? The 100th years anniversary of Trianon and official politics of memory in Hungary. *Rocznik Instytutu Europy Środkowo-Wschodniej*, 18(2), 123–142. <https://doi.org/10.36874/riesw.2020.2.6>
- Fejes, L. (2011, April 7). Egyedül vagyunk a Kárpát-medencében? *Nyelv és Tudomány*. <https://m.nyest.hu/hirek/egyedul-vagyunk-a-karpát-medenceben>
- Gallagher, M., Laver, M., & Mair, P. (2011). *Representative government in modern Europe*. McGraw-Hill.
- Goodson, I. F. & Gill, S. R. (2011). The narrative turn in social research. *Counterpoints*, pp. 386, 17–33.
- Hajdú, Z. (2018). The rebirth of the concept of the Carpathian Basin in Hungarian political language after 1988. In J. Laine, I. Liikanen & J.W. Scott (Eds.), *Post-cold war borders: reframing political space in Eastern Europe* (pp. 207–227). Routledge.

- Horváth, A. & Soós, G. (2015). Pártok és pártrendszer. In A. Körösenyi (Ed.), *A magyar politikai rendszer – negyedszázad után* (pp. 249–278). Osiris Kiadó, MTA TK Politikatudományi Intézet.
- Ilie, C. (2003). Discourse and metadiscourse in parliamentary debates. *Journal of Language and Politics*, 1(2), 269–291. <https://doi.org/10.1075/jlp.2.1.05ili>
- Ilie, C. (2017). Parliamentary debates. In *The Routledge Handbook of Language and Politics* (pp. 309–325). Routledge.
- Indig, B., Sass, B., Simon, E., Mittelholcz, I., Vadász, N., & Makrai, M. (2019). One format to rule them all – The emtsv pipeline for Hungarian. In *Proceedings of the 13th Linguistic Annotation Workshop* (pp. 155–165). Association for Computational Linguistics (ACL).
- Kiss, B. (2019). *A szavakon túl. Politikai kommunikáció Magyarországon, 1990 – 2015*. L'Harmattan Kiadó.
- Kiss, T. (2018). Az anyaország nemzetpolitikai dimenziói és az erdélyi magyarságra gyakorolt hatásai. *Pro Minoritate*, 2018(3), 51–83.
- Lacoste, Y. (1997). *Vive la Nation: Destin d'une idée géopolitique*. Fayard.
- Lesińska, M., & Héjj, D. (2021). Pragmatic trans-border nationalism: A comparative analysis of Poland's and Hungary's policies towards kin-minorities in the twenty-first century. *Ethnopolitics*, 20(1), 53–66. <https://doi.org/10.1080/17449057.2020.1808324>
- Medby, I. & Thornton, P. (2023). More than words: Geopolitics and language. *Area*, 55(1), 2–9. <https://doi.org/10.1111/area.12817>
- Mouffe, C. (2011). *On the political*. Routledge.
- Müller, M. (2008). Reconsidering the concept of discourse for the field of critical geopolitics: Towards discourse as language and practice. *Political Geography*, 27(3), 322–338. <https://doi.org/10.1016/j.polgeo.2007.12.003>
- Nordsieck, W. (2022). *Parties and elections in Europe- Hungary*. Parties and elections in Europe <http://www.parties-and-elections.eu/hungary.html>
- Paasi, A. (2016). Dancing on the graves: Independence, hot/banal nationalism and the mobilization of memory. *Political Geography*, 54, 21–31. <https://doi.org/10.1016/j.polgeo.2015.07.005>
- Pogonyi, Sz. (2017). National Reunification Beyond Borders: Diaspora Politics in Hungary Since 2010. In Sz. Pogonyi (Ed.), *Extra-Territorial Ethnic Politics, Discourses and Identities in Hungary* (pp. 73–123). Springer Nature.
- Roberts, M. E., Stewart, B. M., & Tingley, D. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2). <https://doi.org/10.18637/jss.v091.i02>
- Romsics, G. (2006). Trianon a Házban: A Trianon-fogalom megjelenése és funkciói a pártok diskurzusaiban az első három parlamenti ciklus idején, 1990-2002. In G. Czoch & Cs. Fedinec (Eds.), *Az emlékezet konstrukciói: példák a 19-20. századi magyar és közép-európai történelemből* (pp. 32–52). Teleki László Alapítvány.
- Scott, J. W., & Hajdú, Z. (2022). The Carpathian basin as a 'Hungarian neighbourhood': Imaginative geographies of regional cooperation and national exceptionalism. *Eurasian Geography and Economics*, 63(6), 753–778. <https://doi.org/10.1080/15387216.2022.2082995>

- Szente, Z. (2020). Parlamenti jogunk nyomorúsága. *Fundamentum*, 24(4), 5–19.
- Szücs, G. Z. (2012). A magyar politikai diskurzus változásai 2000 óta. In Zs. Boda & A. Körösenyi (Eds.), *Van irány? Trendek a magyar politikában*. MTA TK PTI – Új mandátum.
- Szücs Z. G. (2015). A magyar politikai gondolkodás nemzetközi horizontja. In A. Körösenyi (Ed.), *A magyar politikai rendszer – negyedszázad után* (pp. 355–375). Osiris – MTA Társadalomtudományi Kutatóközpont PTI.
- Váradi, T., Simon, E., Sass, B., Mittelholcz, I., Novák, A. & Indig, B. (2018). E-magyar – a digital language processing system. In N. Calzolari et al. (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 1307–1312). European Language Resources Association (ELRA).
- Vida A. (2002). Nemzetkonceptiók és státusztörvény: a szomszédos országokban élő magyarokról szóló törvényjavaslat parlamenti vitájának elemzése. *Jel-Kép: Kommunikáció Közvélemény Média*, 21(4), 3–21.
- Waterbury, M. A. (2010). *Between State and Nation: Diaspora Politics and Kin-State Nationalism in Hungary*. Palgrave Macmillan.
- Wodak, R., & Forchtner, B. (Eds.). (2017). *The Routledge handbook of language and politics*. Routledge.
- Wodak, R. (2010). The discursive construction of history – brief considerations. *Mots. Les langages du politique*, (94), 57–65.
- Wodak, R., De Cillia, R., Reisigl, M., & Liebhart, K. (2009). *The discursive construction of national identity* (A. Hirsch & R. Mitten, Trans.; 2nd ed.). Edinburgh University Press.

## Appendix

**Table A1** Distribution by speech type on the full corpus and the corpus filtered by 'Carpathian Basin' keyword

Type of speech	All speeches	Speeches of the Carpathian Basin subcorpus		
	Frequency	Proportion	Frequency	Proportion
Accepted answer by MP	1206	0.71%	2	0.14%
Answer by rapporteur	4626	2.71%	42	2.88%
Comment on a post-agenda speech	608	0.36%	8	0.55%
Comment on a pre-agenda speech	5417	3.18%	94	6.44%
Exposition	2820	1.65%	45	3.08%
Interpellation/question/prompt question	15260	8.95%	65	4.45%
Justification for exceptional procedure	2	0.00%	0	0.00%
Justification of an individual motion	496	0.29%	6	0.41%
Justification of urgency	87	0.05%	0	0.00%
Keynote speech	9311	5.46%	210	14.38%
Opening remarks by rapporteur	913	0.54%	22	1.51%
Oral answer to interpellation	3204	1.88%	15	1.03%
Other type of speech (uncategorized)	59	0.03%	0	0.00%
Post-agenda speech	3256	1.91%	122	8.36%
Pre-agenda speech	6211	3.64%	153	10.48%
Presentation of committee's minority opinion	2559	1.50%	6	0.41%
Presentation of opinion of a committee	6085	3.57%	37	2.53%
Proposal for agenda	2	0.00%	0	0.00%
Question answered	7475	4.38%	36	2.47%
Question for agenda	2018	1.18%	0	0.00%
Recommend for general discussion	1	0.00%	0	0.00%
Rejected answer by MP	2596	2.11%	6	0.41%
Reply to a prompt question by MP	5847	3.43%	9	0.62%
Reply to prompt question by minister	4077	2.39%	7	0.48%
Speech (general category)	38327	22.49%	533	36.51%
Speech on grounds of personal involvement	219	0.13%	0	0.00%
Two-minute speeches	46770	27.44%	107	7.33%
Total	170452	100%	1460	100%

**Table A2** Some representative uses of the term ‘Carpathian Basin’ by topic, period and political orientation

<b>Topic 1, 1998–2010, left-liberal</b>	
‘A Kárpát-medencét természeti viszonyai kiválóan alkalmassá teszik több mezőgazdasági és kertészeti faj vetőmagjának és szaporítóanyagának termesztésére. Ennek megfelelően hazánk világviszonylatban is jelentős szerepet játszik ebben az ágazatban.’	‘The natural conditions of the Carpathian Basin make it an excellent place for the production of seeds and propagating material for many agricultural and horticultural species. Accordingly, our country plays a significant role in this sector worldwide.’
‘[...] ide folyik tulajdonképpen mindenhonnet a Kárpát-medencéből összegyűlő vízmennyiség.’	‘[...] in fact, water from all over the Carpathian Basin flows here.’
<b>Topic 1, 1998–2010, Fidesz</b>	
‘[...] és a korábban kiváló adottságaink nagyon könnyen kedvezőtlenre fordulhatnak itt, a Kárpát-medencében, ahol egyébiránt kiválóak az élelmiszer-előállítás feltételei [...]’	‘[...] and our previously excellent conditions can very easily turn unfavourable here in the Carpathian Basin, where the conditions for food production are excellent [...]’
‘[...] termőföldi adottságaink világviszonylatban is a legjobbak közé sorolhatók, amely után az érdeklődés a Kárpát-medencei tartózkodásunk óta jelentősnek mondható, sok esetben nehézséget okozott az érdeklődők távol tartása is [...]’	‘[...] our farmland is among the best in the world, and since we have been in the Carpathian Basin, interest in it has been considerable, and in many cases it has been difficult to keep people away [...]’
<b>Topic 1, 1998–2010, right-wing nationalist</b>	
‘Az erdőgazdálkodás Magyarországon azóta, tehát 130 év óta tízéves tervek alapján történik. Ennek a 130 év óta folytatott tervszerű munkának az eredménye véleményem szerint a Kárpát-medence erdeinek fennmaradása.’	‘Forest management in Hungary has been based on ten-year plans for 130 years. The result of this 130 years of planned work is, in my opinion, the survival of the forests of the Carpathian Basin.’
‘[...] azt is elmondhatjuk, hogy elsősorban rajtunk múlik az, hogy ezen a veszélyeztetett helyen milyen minőségű életet tudunk magunknak biztosítani; rajtunk és némileg természetesen a szomszédos országokon, mert a jelen politikai helyzetben az egész Kárpát-medence nem tartozik a fennhatóságunk alá.’	‘[...] we can also say that the quality of life in this endangered place is primarily up to us; to us and to some extent, of course, to the neighbouring countries, because in the current political situation the whole Carpathian Basin is not under our jurisdiction.’
<b>Topic 1, 2010–2020, left-liberal</b>	
‘ha a globális átlaghőmérséklet emelkedése eléri a 2 fokot, az a Kárpát-medencében 4 fokos felmelegedést fog eredményezni majd’	‘if the global average temperature increase reaches 2 degrees Celsius, the Carpathian Basin will experience 4 degrees Celsius of warming’
‘Sajnos, a Kárpát-medence és Magyarország is különösen veszélyeztetett.’	‘Unfortunately, the Carpathian Basin and Hungary are particularly at risk.’

Table A2 (continued)

Topic 1, 2010–2020, Fidesz	
‘[...] ha az erőforrásokat nem tudjuk garantáltan ezekhez a helyi közösségekhez, családokhoz, települési közösségekhez juttatni, akkor valószínűleg nagyon súlyos nemzetbiztonsági helyzetbe kerülhet a magyarság itt a Kárpát-medencében.’	‘[...] if we cannot guarantee that the resources will reach these local communities, families and municipal communities, then the Hungarian people here in the Carpathian Basin will probably find themselves in a very serious national security situation.’
‘Az anyaföldben benne van minden kötődésünk a Kárpát-medencei tájhoz, a lakóhelyhez, családtagjainkhoz, mindannyiunkhoz, akik itt élünk, [...]’	‘The motherland contains all our ties to the landscape of the Carpathian Basin, to the place where we live, to our family members, to all of us who live here [...]’
Topic 1, 2010–2020, right-wing nationalist	
‘Az elfogadott Alaptörvényünk már eleve tartalmazza azt a gondolatot, hogy meg kell óvnunk a Kárpát-medence természet adta értékeit [...]’	‘Our adopted Fundamental Law already contains the idea that we must protect the natural values of the Carpathian Basin [...]’
‘[...] arra szeretném kérni a köztársasági elnök urat [...], hogy legyen szíves, támogassa [...] a Kárpát-medence természeti értékeinek a jövő generációk magyarjai számára történő megőrzését.’	‘[...] I would like to ask the President of the Republic [...] to kindly support [...] the preservation of the natural values of the Carpathian Basin for the benefit of future generations of Hungarians.’
Topic 2, 1998–2010, left-liberal	
‘[...] Kárpát-medence népeit – hangsúlyozom, nemcsak a magyarságot, hanem a Kárpát-medence népeit [...]’	‘[...] the peoples of the Carpathian Basin – I stress, not only the Hungarians, but the peoples of the Carpathian Basin [...]’
‘Ez az ország érdeke, ez az itt lakó állampolgárok mindegyikének az érdeke, ez a Kárpát-medencében lakók érdeke, és ez Európa érdeke is.’	‘It is in the interest of the country, it is in the interest of all its citizens, it is in the interest of the people of the Carpathian Basin, and it is in the interest of Europe.’
Topic 2, 1998–2010, Fidesz	
‘A pálinka közismerten a Kárpát-medence egy speciális itala, a magyarsághoz kapcsolódik, [...]’	‘Pálinka is known to be a special drink of the Carpathian Basin, associated with the Hungarian [...]’
‘De én nemcsak a hazai, tehát a határon belüli területről beszélek, hanem a Kárpát-medencében élő magyarság által termékként megjelenő hungarikumokról is, hisz az, hogy a határok elválasztottak bennünket valamilyen oknál fogva, attól még a magyar nemzetség által lakott területeken a hazai terméknek egyfajta hungarikumi védelmet tudni kell biztosítani.’	‘But I am talking not only about the domestic territory, i.e. the territory within the borders, but also about the Hungaricums that are produced by the Hungarian community living in the Carpathian Basin, because the fact that the borders have separated us for some reason does not mean that the Hungarian products in the areas inhabited by the Hungarian ethnic group should not be able to enjoy a kind of Hungaricum protection.’

Table A2 (continued)

Topic 2, 1998–2010, right-wing nationalist	
‘Hisz a Kárpát-medencében olyan nevelést kaptak a családon belül az emberek évszázadokon keresztül, [...]’	‘In the Carpathian Basin, people have been brought up within the family for centuries in a way that [...]’
‘[...] a Kárpát-medencei képernyőkön láthatóvá váljon a magyarság világtelevíziójának első műsora’	‘[...] the first programme of Hungarian world television to be shown on the screens of the Carpathian Basin’
Topic2, 2010–2020, left-liberal	
‘a Kárpát-medencében élő országok’	‘countries in the Carpathian Basin’
‘Láthatjuk, hogy nagyon komoly politikai megosztottság, politikai problémák is szabdálják az itt élő országokat [...] erősíteni kéne az együttműködését akár a Kárpát-medencét nézzük, akár egy nagyobb képet, egy közép-európai térséget az itt élő államoknak, hiszen mondhatjuk, hogy egy sorsközösségben vagyunk [...]’	‘We can see that there are very serious political divisions and political problems among the countries living here [...] we should strengthen cooperation, whether we look at the Carpathian Basin or at the bigger picture, at the Central European region of the states living here, because we can say that we are in a community of destiny [...]’
Topic2, 2010–2020, Fidesz	
‘meghallgassuk, egyáltalán mozgósítsuk az országot, a Kárpát-medencét ebben az ügyben’	‘to listen, to mobilise the country, the Carpathian Basin in this matter’
‘amikor a parlament nyilvánvalóvá tette, hogy nemcsak országhatáron belüli magyar értékekre, hungarikumokra ügyel, [...], hanem a Kárpát-medence, egyáltalán a magyarság értékeit’	‘when the Parliament made it clear that it does not only care about Hungarian values and Hungaricums within the national borders [...], but also about the values of the Carpathian Basin, of the Hungarian people in general.’
Topic2, 2010–2020, right-wing nationalist	
‘A törvényjavaslat szerint az Országgyűlés többek között megállapítja, hogy [...] a Kárpát-medence ember alkotta és természet adta értékeit átfogó értéktárban kell összesíteni.’	‘According to the bill, the Parliament states, among other things, that [...] the man-made and natural values of the Carpathian Basin shall be collected within a comprehensive inventory of values.’
‘számunkra soha nem külföld Kárpát-medence megszállt magyar területe, bármelyik magyar területe, de formáljogilag külföld’	‘is never a foreign country for us, the occupied Hungarian territory of the Carpathian Basin, any Hungarian territory, but formally, legally a foreign country’
Topic3, 1998–2010, left-liberal	
‘ez a Kárpát-medencei magyarságra terjedne ki’	‘it would cover the Hungarian population of the Carpathian Basin’
‘fokozott a felelősségünk, hogy az Európai Unió bővítése mellett hogyan gondoskodunk azokról a magyarokról, akik Kárpátján és a Vajdaságban élnek’	‘we have an increased responsibility to take care of the Hungarians living in the Carpathian Basin and Vojvodina beside the enlargement of the European Union.’



Table A2 (continued)

Topic3, 1998–2010, Fidesz	
‘nekünk, akik a Kárpát-medencében akarunk és tudunk gondolkodni, létérdekünk lenne, hogy minden jogot, amely ma állampolgársághoz kötődik, áttörjünk.’	‘it would be in the vital interest of us who want to and can think in the Carpathian Basin to break through all the rights that are now linked to citizenship.’
Topic3, 1998–2010, right-wing nationalist	
‘Nem az ország tévedt rossz útra – ahogy a miniszterelnök úr fogalmaz -, ez az ország egyrészt nem kel útra, ez itt van, ezer éve itt van a Kárpát-medencében, Európában, a kormány tévedt rossz útra, az önök kormánya tévedt rossz útra, az ország pedig csak elszenvedője az önök szélhámosságának’	‘The country has not gone astray – as the Prime Minister puts it – on the one hand, this country is not going astray, it is here, it has been here for a thousand years in the Carpathian Basin, in Europe, the government has gone astray, your government has gone astray, and the country is just a victim of your deceit.’
Topic3, 2010–2020, left-liberal	
‘a környező országok, a Kárpát-medencében lévő országok’	‘neighbouring countries, countries in the Carpathian Basin’
Topic3, 2010–2020, Fidesz	
‘nekünk, képviselőknek, mindnyájunknak dolgozni kell, hogy a Kárpát-medence ezer esztendő múlva is magyar szótól legyen hangos, a magyar legyen akkor is a legnagyobb nemzet a Kárpát-medencében, és mi irányítsuk ezt a medencét’	‘we MPs, all of us, must work to ensure that the Carpathian Basin will still be full of the Hungarian word a thousand years from now, that Hungarians will still be the largest nation in the Carpathian Basin, and that we will still be in charge of this basin’
‘a nyílt társadalom elsőszámú lobbistájával van dolgunk, akit teljesen hidegen hagy, hogy itt a Kárpát-medencében száz év múlva magyarok élnek-e vagy sem.’	‘we are dealing with the number one lobbyist of open society, who is completely indifferent to whether or not there will be Hungarians living here in the Carpathian Basin in a hundred years’ time.’
Topic3, 2010–2020, right-wing nationalist	
‘a kormánynak törekednie kell a Kárpát-medencei magyarság természetes egységének helyreállítására’	‘the government should strive to restore the natural unity of the Carpathian Basin Hungarians’
Topic 4, 1998–2010, left-liberal	
‘Nem egyedül éltük meg nehéz történelmünket, az évezredet, s ma sem egyedül élünk a Kárpát-medencében.’	‘We have not lived our difficult history, this millenium alone, and we do not live alone in the Carpathian Basin today.’
Topic 4, 1998–2010, Fidesz	
‘Mindez alapja lehet Magyarországnak, a Kárpát-medence, a bíboros szóhasználatával élve: a történelmi haza erkölcsi, szellemi megújulásának is.’	‘All this can be the basis for the moral and spiritual renewal of Hungary, of the Carpathian Basin, to use the Cardinal’s words, of the historic homeland.’



Table A2 (continued)

‘[...] hogy megszólítsam az Országgyűlést, az ország népét, a Kárpát-medence népét, a magyar nemzetet szerte a glóbuszon’	‘[...] to address the National Assembly, the people of the country, the people of the Carpathian Basin, the Hungarian nation all over the globe’
<b>Topic 4, 1998–2010, right-wing nationalist</b>	
‘mi a titka annak, hogy ezer olyan nehéz esztendő után ez a maroknyi nép itt élhet a Kárpát-medencében’	‘what is the secret that after a thousand years of hardship this handful of people can live here in the Carpathian Basin’
<b>Topic 4, 2010–2020, left-liberal</b>	
‘dicséretes, hogy a nemzeti összetartozás fogalomköréből továbblépve egy tágabb integráció fontosságáig tágítja a kört, jelesül a Kárpát-medencei népek együtt élésének, egymásra utaltságának fontosságát hangsúlyozza’	‘it is commendable that, moving on from the concept of national belonging, it broadens the scope to the importance of a broader integration, namely it stresses the importance of the Carpathian Basin peoples living together and their interdependence’
<b>Topic 4, 2010–2020, Fidesz</b>	
‘A magyar állam a Kárpát-medence államaiban és a diaszpórában élő magyar közösségeket az egységes magyar nemzet részének tekinti’	‘The Hungarian state considers the Hungarian communities living in the Carpathian Basin states and in the diaspora as part of the united Hungarian nation’
<b>Topic 4, 2010–2020, right-wing nationalist</b>	
‘abba, ami 1920-ban történt, a Kárpát-medence földrajzi, gazdasági, politikai és kulturális egységének szétszaggatásába, Szent István országának csonkolásába ne törődjenek bele’	‘not to put up with what happened in 1920, in the dismemberment of the geographical, economic, political and cultural unity of the Carpathian Basin, in the mutilation of the country of St Stephen’
<b>Topic 5, 1998–2010, left-liberal</b>	
‘készül egy agrárfejlesztési program, ‘Híd a harmadik évezredbe,’ amely az egész Kárpát-medencére ki fog terjedni’	‘an agricultural development programme, ‘Bridge to the third millennium,’ is being prepared, which will cover the whole Carpathian Basin’
<b>Topic 5, 1998–2010, Fidesz</b>	
‘[...] a Kárpát-medence területén is megszorodtak a szélsőséges természeti jelenségek. Az általuk okozott kár három esztendő leforgása alatt több mint 150 milliárd forint veszteséget okozott a nemzetgazdaság számára.’	‘[...] extreme natural phenomena have also increased in the Carpathian Basin. The damage caused by these events has resulted in losses of more than 150 billion forints for the national economy in three years’
<b>Topic 5, 1998–2010, right-wing nationalist</b>	
‘A Magyar Demokrata Fórum célja, hogy ebben az évezredben is egy erős, életképes, nemzeti hagyományait, kultúráját ápoló és fejlesztő magyar nemzet éljen a Kárpát-medencében. [...] Ennek a politikának az egyik legfontosabb eleme a népesedéspolitika, családpolitika folytatása.’	‘The goal of the Hungarian Democratic Forum is to ensure that a strong, viable Hungarian nation that cherishes and develops its national traditions and culture will continue to live in the Carpathian Basin in this millennium. [...] One of the most important elements of this policy is the continuation of population and family policy.’

Table A2 (continued)

Topic 5, 2010-2020, left-liberal	
‘A kormány újra feléleszti az MFB-n belül a határon túli, Kárpát-medencei térségbe irányuló tőkebefektetéseket, ami öt-hat éve sajnos elhalt.’	‘The government is reviving capital investment in the cross-border Carpathian Basin region within the MFB, which unfortunately died out five or six years ago.’
Topic 5, 2010-2020, Fidesz	
‘Különösen, ha Romániára gondolunk, nemcsak Bukarestből, hanem Nagyváradról is lehet már New Yorkba repülni. Igen-igen nagy fejlesztések indultak meg a Kárpát-medencében és a közép-kelet-európai régióban.’	‘Especially if you think of Romania, you can now fly to New York not only from Bucharest, but also from Oradea. Yes, there are very, very big developments in the Carpathian Basin and in the Central and Eastern European region.’
Topic 5, 2010-2020, right-wing nationalist	
‘Nem nehéz belátni, ha bármilyen Kárpát-medencét illető olyan tervünk van, amely egy Kárpát-medencei egységes gazdasági övezetet magyar dominancia mellett kíván kialakítani, [...]’	‘It’s not difficult to see that if we have any plans for the Carpathian Basin that would create a single economic zone in the Carpathian Basin with Hungarian dominance, [...]’
Topic 6, 1998-2010, left-liberal	
‘Szeretném hozzátenni az anyagnak azt az elemét is, amely a régióban is gondolkodik, és azt szorgalmazza, hogy itt a Kárpát-medence vagy Közép-Kelet-Európa térségében is érdemes lenne egy közös stratégia mentén [...]’	‘I would also like to add the element of the document that also thinks in terms of the region, and urges that it would be worthwhile to have a common strategy for the Carpathian Basin or Central-Eastern Europe [...]’
Topic 6, 1998-2010, Fidesz	
‘Számunkra kizárólag egységes magyar nemzeti kultúra létezik, szülessen az a kultúra bárhol a Kárpát-medencében vagy az emigráció körében.’	‘For us, there is only a single Hungarian national culture, whether it is born anywhere in the Carpathian Basin or among the emigrants.’
Topic 6, 1998-2010, right-wing nationalist	
‘Fontos lenne ugyanakkor a régiók műemlékeit együttesen kezelni, egy átfogó rendezési tervet kidolgozni rájuk az EU keretében is, hiszen sok kárpát-medencei épület együttes rendbehozása is beletartozna ebbe a körbe.’	‘At the same time, it would be important to treat the monuments of the regions together, to develop a comprehensive management plan for them within the framework of the EU, as the restoration of many Carpathian Basin monuments would also be included in this scope.’
Topic 6, 2010-2020, left-liberal	
‘az északkeleti Kárpát-medence fatemplomai’	‘the wooden churches of the north-eastern Carpathian Basin’
Topic 6, 2010-2020, Fidesz	
‘Kárpát-medencei borásztársadalom’	‘winemaking community in the Carpathian Basin’

Table A2 (continued)

Topic 6, 2010–2020, right-wing nationalist	
‘olyan ország vagyunk, akiknek nemzeti kulturális örökségének jelentős része határon túlra került, így nekünk, az anyaországnak az is feladatunk, hogy a Kárpát-medence egységes és oszthatatlan magyar nemzeti örökségére vigyázzunk’	‘we are a country where a significant part of our national cultural heritage has been transferred across borders, so it is also our duty as the motherland to take care of the united and indivisible Hungarian national heritage of the Carpathian Basin’
Topic 7 1998–2010, left-liberal	
‘Alapvető a Kárpát-medencében élő magyarok számára, hogy folytatódik és reméljük, befejeződik az a folyamat, amely újraegyesíti a szétszakított magyarságot az európai integráció révén a határok megváltoztatása nélkül’	‘It is fundamental for the Hungarians living in the Carpathian Basin that the process of reuniting the divided Hungarians through European integration without changing borders will continue and, we hope, be completed.’
Topic 7 1998–2010, Fidesz	
‘a polgári koalíció kormányának célkitűzéseiben is kiemelt feladatként jelent meg a Kárpát-medencei és a nyugati magyarság rendszeresebb, szorosabb, úgymond intézményesített kapcsolattartása az anyaországgal.’	‘in the objectives of the government of the civic coalition, a more systematic, closer, so to speak institutionalised, relations between the Carpathian Basin and the Hungarians of the West and the motherland were also a priority.’
Topic 7 1998–2010, right-wing nationalist	
‘a Kárpát-medencében élő magyarságnak egyetlen történelmi realitása az újraegyesítés kapcsán az európai uniós csatlakozás.’	‘for the Hungarian people living in the Carpathian Basin the only historical reality of the reunification is European Union accession.’
Topic 7 2010–2020, left-liberal	
‘megadunk minden támogatást a kárpátaljai és az egész Kárpát-medencei magyar oktatás megmaradásáért a magyar kormánynak ‘	‘we give all support to the Hungarian government for the survival of Hungarian education in Transcarpathia and the entire Carpathian Basin’
Topic 7 2010–2020, Fidesz	
‘A Kárpát-medencei autonómiapolitika történetében [...]’	‘In the history of autonomy policy in the Carpathian Basin [...]’
Topic 7 2010–2020, right-wing nationalist	
‘Amikor tehát határon túli magyarokról beszélünk, akkor azért mindig tegyük zárójelbe ezt a jelzőt; beszéljünk magyarságról, beszéljünk Kárpát-medencei és a világban élő magyarságról, beszéljünk legfeljebb elcsatolt vagy elszakított testvéreinkről, de azt gondolom, hogy nem lehet beállni abba az irányba, amit többek között például az a nemzeti érdekeket feladó politika testesített meg [...]’	‘So when we talk about Hungarians beyond the borders, we should always put this adjective in brackets; let’s talk about Hungarians, let’s talk about Hungarians living in the Carpathian Basin and in the world, let’s talk about our annexed or separated brothers and sisters, but I think that we cannot go in the direction of what, among others, the policy of abandoning national interests has embodied [...]’

Table A2 (continued)

Topic 8 1998–2010, left-liberal	
‘[...] a Kárpát-medencében élő magyar fiatalok számára a két legfontosabbnak tartott érték a család, illetve a saját jövő’	‘[...] for young Hungarians living in the Carpathian Basin, the two most important values are family and their own future.’
Topic 8 1998–2010, Fidesz	
‘[...] amely lehetőséget biztosít a folyamatos párbeszédre a Kárpát-medence minden tájáról származó magyar fiataloknak’	‘[...] which provides an opportunity for continuous dialogue for young Hungarians from all over the Carpathian Basin’
Topic 8 1998–2010, right-wing nationalist	
‘a legnagyobb probléma a nemzetstratégia kérdésében, a közös Kárpát-medencei nemzetstratégiánk összefüggésében a bizalmatlanság’	‘the biggest problem in the issue of national strategy, in the context of our common national strategy for the Carpathian Basin, is mistrust’
Topic 8 2010–2020, left-liberal	
‘a Kárpát-medencében a határon túli magyarsággal kapcsolatos ismeretek bővülése osztálykirándulások kapcsán meg tudjon valósulni [...]’	‘[...] which would make it possible to increase the knowledge about Hungarians living beyond the borders in the Carpathian Basin through class excursions [...]’
Topic 8 2010–2020, Fidesz	
‘[...] megkezdődött az egyes Kárpát-medencei magyar régiók fejlesztését célzó gazdaságfejlesztési tervek támogatása’	‘[...] support for economic development plans aimed at the development of certain Hungarian regions in the Carpathian Basin has started’
Topic 8 2010–2020, right-wing nationalist	
‘A Kárpát-medencében a történelmi egyházaink egy élő közösséget és működő hálózatot jelentenek’	‘Our historic churches in the Carpathian Basin are a living community and a functioning network’

Table A3 Change in topic prevalence by ideological bloc

Term		Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
<b>FIDESZ</b>									
1	Mean	0.18	0.05	0.25	0.14	0.11	0.03	0.20	0.04
	Standard error of mean	0.03	0.02	0.03	0.02	0.02	0.01	0.02	0.01
2	Mean	0.14	0.07	0.13	0.19	0.07	0.11	0.17	0.12
	Standard error of mean	0.02	0.01	0.01	0.02	0.01	0.01	0.01	0.01
Total	Mean	0.15	0.07	0.16	0.17	0.08	0.09	0.18	0.10
	Standard error of mean	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
<b>RIGHT-WING NATIONALIST</b>									
1	Mean	0.20	0.07	0.08	0.19	0.13	0.05	0.26	0.03
	Standard error of mean	0.03	0.02	0.01	0.02	0.02	0.01	0.03	0.01
2	Mean	0.11	0.10	0.18	0.16	0.12	0.07	0.16	0.10
	Standard error of mean	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Total	Mean	0.14	0.09	0.16	0.17	0.12	0.06	0.19	0.08
	Standard error of mean	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
<b>LEFT-LIBERAL</b>									
1	Mean	0.06	0.04	0.16	0.20	0.10	0.09	0.32	0.04
	Standard error of mean	0.02	0.01	0.02	0.03	0.02	0.02	0.03	0.01
2	Mean	0.25	0.17	0.15	0.12	0.05	0.06	0.11	0.09
	Standard error of mean	0.03	0.03	0.02	0.02	0.01	0.02	0.02	0.02
Total	Mean	0.16	0.10	0.15	0.16	0.07	0.08	0.21	0.06
	Standard error of mean	0.02	0.01	0.02	0.02	0.01	0.01	0.02	0.01

Table A4 Distribution of topics by type of speech (prevalence with standard error).

Type of speech	Number of speeches	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
Presentation of committee's minority opinion	N=6	0.30	0.01	0.16	0.00	0.33	0.00	0.12	0.07
		0.19	0.01	0.14	0.00	0.18	0.00	0.10	0.06
Answer by rapporteur	N=42	0.12	0.20	0.30	0.08	0.03	0.08	0.16	0.03
		0.04	0.05	0.05	0.03	0.01	0.03	0.05	0.02
Exposition	N=45	0.14	0.08	0.07	0.15	0.10	0.10	0.24	0.12
		0.05	0.03	0.03	0.05	0.04	0.04	0.05	0.04
Presentation of opinion of a committee	N=37	0.06	0.05	0.06	0.06	0.14	0.07	0.25	0.30
		0.03	0.03	0.02	0.03	0.05	0.04	0.06	0.07
Reply to prompt question by MP	N=9	0.14	0.17	0.16	0.14	0.08	0.01	0.27	0.03
		0.11	0.09	0.08	0.09	0.07	0.00	0.10	0.03
Reply to prompt question by minister	N=7	0.15	0.01	0.31	0.11	0.04	0.18	0.09	0.11
		0.13	0.01	0.12	0.07	0.02	0.13	0.05	0.10
Interpellation / question / prompt question	N=65	0.30	0.04	0.04	0.03	0.08	0.11	0.21	0.18
		0.05	0.02	0.01	0.01	0.03	0.03	0.04	0.04
Opening remarks by rapporteur	N=22	0.14	0.14	0.05	0.18	0.01	0.22	0.14	0.13
		0.07	0.07	0.04	0.07	0.00	0.09	0.06	0.07
Speech (general category)	N=533	0.15	0.10	0.18	0.16	0.12	0.07	0.15	0.08
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Oral answer to interpellation	N=15	0.08	0.05	0.18	0.03	0.13	0.10	0.29	0.15
		0.06	0.05	0.08	0.03	0.07	0.06	0.09	0.07
Accepted answer by MP	N=2	0.00	0.01	0.05	0.01	0.01	0.01	0.36	0.54
		0.00	0.00	0.02	0.01	0.00	0.00	0.33	0.32
Rejected answer by MP	N=6	0.14	0.06	0.51	0.06	0.05	0.01	0.13	0.03
		0.09	0.04	0.10	0.04	0.03	0.00	0.07	0.03
Question answered (1)	N=1	0.00	0.00	0.07	0.00	0.00	0.01	0.06	0.84
		.	.	.	.	.	.	.	.

Table A4 (continued)

Type of speech	Number of speeches	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8
Question answered (2)	N=35	0.17	0.08	0.16	0.06	0.09	0.10	0.19	0.15
		0.05	0.03	0.05	0.03	0.04	0.04	0.05	0.04
Two-minute speeches	N=107	0.13	0.13	0.34	0.11	0.10	0.03	0.13	0.04
		0.03	0.03	0.03	0.02	0.02	0.01	0.02	0.01
Pre-agenda speech	N=153	0.14	0.04	0.10	0.31	0.04	0.05	0.28	0.04
		0.03	0.01	0.02	0.03	0.01	0.01	0.03	0.01
Comment on pre-agenda speech	N=94	0.14	0.04	0.19	0.15	0.08	0.03	0.33	0.03
		0.03	0.01	0.03	0.03	0.02	0.01	0.04	0.01
Post-agenda speech	N=122	0.13	0.05	0.04	0.41	0.12	0.08	0.15	0.01
		0.03	0.02	0.01	0.04	0.03	0.02	0.03	0.00
Comment on post-agenda speech	N=8	0.01	0.01	0.09	0.21	0.02	0.16	0.48	0.04
		0.00	0.00	0.04	0.11	0.01	0.11	0.15	0.02
Keynote speech	N=210	0.15	0.09	0.14	0.09	0.10	0.12	0.18	0.13
		0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Justification of individual motion	N=6	0.06	0.09	0.17	0.10	0.14	0.01	0.40	0.03
		0.06	0.07	0.07	0.04	0.12	0.00	0.13	0.02
Total	N=1525	0.14	0.08	0.16	0.17	0.10	0.07	0.19	0.08
		0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

## Online Supplement

The interactive visualization below allows users to explore the results of our model by selecting various parameters. The figure includes language as one of its parameters: terms can be displayed in English as well in the online version. The chart was created based on the results of the 'STM' R-Package. The figure indicates differences in the vocabulary of each topic and illustrates variances in vocabulary across selected political blocs. Line lengths reflect the extent to which words favor one bloc over the other (based on probability of use). Dots are sized proportional to their use within the plotted topic. The bigger the dot, the more emphasized the word. This analytical tool pertains to the substantive content of each topic. While the semantic differences in the three subcorpora may not be consistent for every topic, both minor and major shifts in emphasis equally offer opportunities for interpretation.

[https://public.tableau.com/app/profile/eszterkatona/viz/discourse\\_carpathian\\_basin/Dashboard1?publish=yes](https://public.tableau.com/app/profile/eszterkatona/viz/discourse_carpathian_basin/Dashboard1?publish=yes)



---

\* [\[tury.angyal.emese@tatk.elte.hu\]](mailto:tury.angyal.emese@tatk.elte.hu) (Department of Statistics, Faculty of Social Sciences, Eötvös Loránd University Doctoral School of Sociology and Communications, Corvinus University of Budapest)

\*\* [\[laszlo.lorincz@uni-corvinus.hu\]](mailto:laszlo.lorincz@uni-corvinus.hu) (ANETI Lab, Institute for Data Science and Informatics & Corvinus Institute for Advanced Studies, Corvinus University of Budapest ANETI Lab, Institute of Economics, Centre for Economic and Regional Studies)

### Abstract

Social media plays a crucial role in online political campaigns as political parties can reach, inform, and mobilize voters through these platforms. Political campaigns share information on social media to mobilize support, and prior research shows that sharing content on social media correlates with the offline popularity of political parties. In this paper, we model the spread of political content on the internet. We start by exploring popularity and sharing behavior related to posts by Hungarian politicians on Facebook. We utilize this analysis to build an agent-based model. Within this, we test how echo chambers, homophily, and network structure affect the number of shares that contribute to information diffusion on social media. Our simulation compares spreading in different network structures and shows that preferential attachment models are not the most efficient for fostering diffusion in networks with relatively low density or when a filtering mechanism is present. Our model confirms that homophily generally has a positive effect on diffusion, especially within echo chambers. Echo chambers enhance the diffusion of political news with a limited potential audience. Furthermore, the results of our agent-based simulation indicate that homophily and echo chambers can significantly influence the spread of political content on social media, with echo chambers particularly enhancing diffusion in networks where overall diffusion is low.

**Keywords:** social media, political participation, agent-based model

## 1 Introduction

Political activity is any activity that is intended to or has the consequence of affecting, either directly or indirectly, government action (Verba et al., 1995). It can occur offline, in traditional forms – participating in demonstrations, contacting members of the government, signing a petition, etc. – or, as has become more common in the last few decades, via online platforms on social media sites.

Political campaigns utilize social media sites to engage voters and aim to mobilize them to share political content. For political parties, this seems to be an effective tactic for

---

reaching their ultimate goal – winning elections – as recent research has shown that the number of shares on social media sites may correlate with the offline popularity of a political party (Bene, 2018).

The use of social media during political campaigns has been significant since 2006; political candidates have been using Facebook for political campaign purposes (MacWilliams, 2015), and it has become increasingly frequently used for the organization of political demonstrations (Koltai & Stefkovics, 2018). Magin et al. (2017) argues that political campaigning has three main functions: to disseminate information, to facilitate dialogue between politicians and voters, and to mobilize support.

Prior research has mainly focused on the mobilization aspect and investigated the relationship between online and offline participation (Feezell et al., 2016; Groshek & Dimitrova, 2011; Oser et al., 2013; Theocharis & Lowe, 2016). This paper understands mobilization as sharing content (Klinger, 2013; Magin et al., 2017) and focuses on information dissemination. The dynamics of information dissemination are modeled as the velocity of the spread of political content in various network scenarios. We analyze how characteristics of social media, such as the effects of reactions and comments to a post, network structure, political homophily, and algorithms that amplify echo chambers, may influence online information spread. For this purpose, we introduce an agent-based model based on observations of actual Facebook posts. The results of the model reveal several key findings about the role of network characteristics. In our analysis, small-world networks outperformed preferential attachment networks (Albert & Barabási, 2002) in terms of the number of shares, particularly when network density was low or constrained by filtering algorithms. This is because sharing is influenced by users' political alignment, which can hinder diffusion in highly centralized networks. Furthermore, in the model, homophily has varied effects depending on the network setting. In more connected networks with similar interests, posts spread extensively within that group but struggled to break out of it. Interestingly, in scenarios where diffusion was limited by content-filtering algorithms, homophily enhanced the spread of news, especially in small-world networks. Echo chambers also play a significant role in enhancing diffusion, particularly for news with a limited potential audience. When echo chambers are present, homophily positively affects diffusion by creating paths through which politically interested agents can be reached, even if they are distant in the network.

## 2 Theoretical background

Individuals can benefit from social connections, as they can access and use the resources of other individuals through them. Regarding politicians, this benefit translates into the opportunity to reach otherwise unrecognized people, thus lowering campaign costs (Valenzuela et al., 2018). This benefit, however, depends on whether online activities impact offline political outcomes.

In the light of previous empirical studies, the existence of this link is not evident. While some studies have demonstrated the tangible impact of social media on political participation, like on voting (Dimitrova et al., 2014; Holt et al., 2013), others have found that reliance on social networking sites had no effect on political participation, although

it was related to civic participation (Zhang et al., 2010). Further positive findings include Skoric and Zhu (2016), using Singaporean data, while negative ones include Groshek & Dimitrova (2011), who found no significant impact of social media use on vote intention in the 2008 US presidential election and Theocharis & Lowe (2016), who demonstrated the negative effect (substitution) between online and offline political participation using an experiment on Facebook. To synthesize this mixed evidence, we can rely on a meta-analysis conducted by Boulianne (2020), who investigated this issue using over 300 studies from the past twenty years. The study revealed that despite significant cross-country variation, a positive relationship between online activity and offline political participation exists.

A key mechanism linking online and offline political activity is sharing political content. According to Magin et al. (2017), political campaigning has three main goals: to disseminate information, to facilitate dialogue between politicians and voters, and to mobilize support. Sharing content integrates voters into the campaign: it is a common, low-threshold, but potentially very effective mass-centered form of mobilization. Thus, sharing acts as the key micro-level link between online and offline forms of political participation on social media.

We know that social media connections emerge from distinct real-life social network structures (Vepsäläinen et al., 2017). However, the main attribute of these platforms is to create and display connections with others on a platform via a semi-public or public profile (Boyd & Ellison, 2007). Therefore, network structures can be substantially different across different media platforms and according to the different facets of the connections that are maintained offline. As news shared on social media may spread differently depending on the network structure (Pegoretti et al., 2012), the network structure behind the social media platform is a key factor in the analysis.

The increasing unavailability of large contemporary social media for research purposes, starting with Facebook and followed by X (ex-Twitter), however, restricts the current analysis of the global network structure of these platforms. Still, evidence from earlier studies suggests that the network structure of online social networks combines several features of basic models. Degree distributions in online social networks are definitely unequal, with a few people having many connections and many only a few. However, the former are not as unequal as preferential attachment (PA) models would suggest; rather, the best of the latter involve a hybrid process of random and preferential attachment (Corten, 2012). Further, they show a significant positive degree of assortativity (Corten, 2012; Ugander et al., 2011), which does not follow either from the random or the preferential attachment model (c.f. Barabási & Pósvai Ch. 7). They also involve significant clustering (Corten, 2012; Ugander et al., 2011), which again does not follow from the above models, but is a property of the small world model. A related property is that they exhibit a hierarchical structure in similarity, but two random people can still reach each other over a very small distance (Watts et al., 2012), otherwise known as the six degrees of separation.

Therefore, simulations usually also apply these types of networks to model social network sites: namely, preferential attachment networks, random (a variant of Erdős-Rényi model) networks, and small-world networks (Chan 2019, Jiang & Jiang, 2014).

The impact of network structure has been analyzed using different opinion dynamics and diffusion models. Pegoretti et al. (2012) found that information diffusion is faster in

small-world networks than in random networks when information is not perfect, meaning that information is not known to all members of the network equally but spreads through the ‘demonstration effect,’ i.e., via contacting each other (broadcasting or marketing). Centola (2010) also found that the small world network involves ties that bridge long distances, and the former propagate/diffuse information in an experimental setting faster than in a lattice-structured network. In contrast, random networks perform faster when diffusing an innovation with an equal chance for each agent, who all have an idiosyncratic willingness to adopt (Pegoretti et al., 2012). Korkmaz et al. (2019) found that diffusion was generally faster on scale-free networks than on random networks.

We can assess different features of ‘real’ social networks using these basic network models, albeit separately. Basically, they all generate relatively short distances (compared to the lattice approach). The preferential attachment mechanism creates hubs, while the small world network creates local cohesion (clustering). However, echo chambers are a central aspect of online politics that we cannot analyze with these network properties.

Echo chambers are defined as clusters formed by users with homogeneous content production and diffusion, in which one’s beliefs are reinforced due to repeated interactions with individuals sharing the same points of view (Cota et al., 2019). Selective exposure (homophily) and confirmation bias are key mechanisms contributing to the formation of echo chambers (Quattrociocchi et al., 2021).

In networks, homophily is defined as the inclination of people to interact more with others with similar characteristics rather than with people with different ones. This emerges along two key social dimensions: status and values (Lazarsfeld & Merton, 1954; McPherson et al., 2001). From our perspective, the relevant dimension is value homophily, that is, whether people who are more aligned with a political opinion and, therefore, more likely to share it are more likely to be connected in the network.

A further element that may amplify echo chambers is the algorithms used by social media. Personalized recommender algorithms are routinely used by e-commerce and social media to filter content that fits the preferences of the user (Ge et al., 2020). Recommending friends on social media itself contributes to echo chambers if homophily is present (Cinus et al., 2022). In addition, the presence of content filtering according to the preferences of the user contributes to the positive feedback loop of echo chambers (Jiang et al., 2021).

As basic network models themselves do not create echo chambers, in order to analyze them in the model, we add these two features, homophily, and content filtering, to the model as additional mechanisms.

Previous studies have found that homophily significantly influences diffusion. Aral et al. (2013) found that the adoption of a new (instant messaging) service significantly decreased in real networks compared to reshuffled networks from which homophily was eliminated; thus, homophily decreased diffusion. Korkmaz et al. (2019) have shown that both homophily and heterophily are better than random assignment in terms of the speed and size of cascades in an observed network and a scale-free network model, but in random networks, homophily promoted diffusion. Simulation models have shown that social influence in opinion dynamics and echo chambers in the case of controversial issues leads to the polarization of opinions instead of developing a consensus and the segregation of the network into several separated communities (Baumann et al., 2020; Li & Tang, 2015).

Such dynamics of polarization have also been observed empirically on social networks (Del Vicario et al., 2016; Li & Tang, 2015). From the point of view of news sharing, such a polarized outcome may correspond to the limited diffusion of news, wherein the news reaches only that cluster of users who initially had favorable attitudes. Concerning the relationship between the diffusion of the news and recommender algorithms, Quattrociocchi et al. (2021) found greater segregation in news consumption on Facebook than in Reddit and larger biases in information diffusion due to the clusters on social media based on content curating algorithms that are not tweakable by users (Facebook, Twitter) in contrast to other platforms, e.g., Reddit. In the case of Facebook, they found that the user's attitude ('leaning') affects who the final recipients of the information are, thus increasing the polarization in information diffusion. Considering news sharing, therefore, we expect that a content filtering algorithm (based on preferences) itself will limit the diffusion of news sharing and that the negative impact of homophily and preference-based filtering algorithms may amplify each other.

To sum up, previous research has produced divergent conclusions about the impact of network structures on diffusion. However, the modeling assumptions were also heterogeneous. Therefore, our key research question is: How do the above-described results apply to the case of sharing political news online? Which network structures will be more efficient in our case? While we do not have a clear hypothesis concerning network structure based on the previous literature, with a focus on news sharing, it is anticipated that a content-filtering algorithm based on user preferences could potentially restrict the diffusion of news sharing. Additionally, we assume that the negative influence of homophily (individuals tend to connect with like-minded people) and preference-based filtering algorithms may amplify each other, leading to further limitations on the dissemination of political news.

### 3 Data and methods

#### 3.1 Methods

To model information spread on networks, we use agent-based modeling (ABM). ABM, also named individual-based modeling, is a method of modeling dynamics in complex systems that is often used to study the emergence of macro-level phenomena from individual, micro-level interactions in the social sciences. In an agent-based simulation, agents are autonomous, interactive individuals who keep evolving by monitoring their neighbors' state through stochasticity or, as in sophisticated agent-based modeling scenarios, through artificial intelligence approaches (Helbing, 2012). ABM is also an appropriate and widely used methodology for modeling complex phenomena in various network architectures (Ylikoski, 2014), such as information diffusion and the relationship between online and offline political activity.

Agent-based models have the advantage of a high degree of flexibility; they are capable of including various simple or complex mechanisms of interaction across agents. This also means that the model specification needs external inputs and some theoretical guidance on building the interactions. For this purpose, we build on pre-existing models of diffu-

sion and social influence. However, in order to get insights on how to apply them to our specific case, spreading political news, we also turn to an empirical examination of social media.

### 3.2 Data Analysis

To create an empirical starting point for our agent-based model, we analyze Facebook data about the political activity of Hungarian social media users. Of the two most popular current social media platforms that are used in politics, namely X (ex-Twitter) and Facebook, we chose Facebook because, in Hungary, Twitter is way less popular than Facebook. Bene and Somodi (2018) have shown that Hungarian politicians are typically available on Facebook.

In addition, although data was more readily available for analyzing Twitter usage, through the dedicated software tool Crowdtangle, Facebook's data was also made available for limited research purposes. Crowdtangle is a Facebook-owned tool that tracks interactions on public content from Facebook pages, groups, verified profiles, Instagram accounts, and subreddits. It does not include paid ads unless those ads began as organic non-paid posts that were subsequently 'boosted' using Facebook's advertising tools. It does not include activity by private accounts or posts made visible only to specific groups of followers, either.

For the analysis, the data collection period ran from January 21-22 (Friday-Saturday), 2021, and targeted seven Hungarian political parties and their leaders. Some of the parties had joint leadership; in these cases, pages for both leaders were included. The data covers every post these parties and party leaders created in the given period. The data also includes the number of reactions and shares associated with each post in consecutive time-steps. The last timestep is the 74th, which marks that at least 20 days have passed since the original post.

During the campaign, several parties from the opposition formed an electoral alliance and have been campaigning together since the end of 2020. The electoral alliance was made up of six political parties: Demokratikus Koalíció – DK (Democratic Coalition - DK); LMP – Magyarország Zöld Pártja (LMP – Hungary's Green Party); Jobbik Magyarországért Mozgalom (Jobbik - Movement for a Better Hungary); Momentum Mozgalom (Momentum Movement); Magyar Szocialista Párt – MSZP (Hungarian Socialist Party – MSZP); and Párbeszéd Magyarországért (Párbeszéd – Dialogue for Hungary). Polls measured this alliance's popularity as increasing, closing in on the incumbents' popularity (Közüvéleménykutatók.hu).<sup>1</sup> However, this electoral alliance was a loose formation, and the participating parties decided to hold a primary in late 2021 to select a final candidate to run for the office of the Hungarian Prime Minister in 2022. Thus, parties from the opposition also campaigned against each other throughout most of 2021 and only showed a united front after the primaries in the autumn of 2021.

---

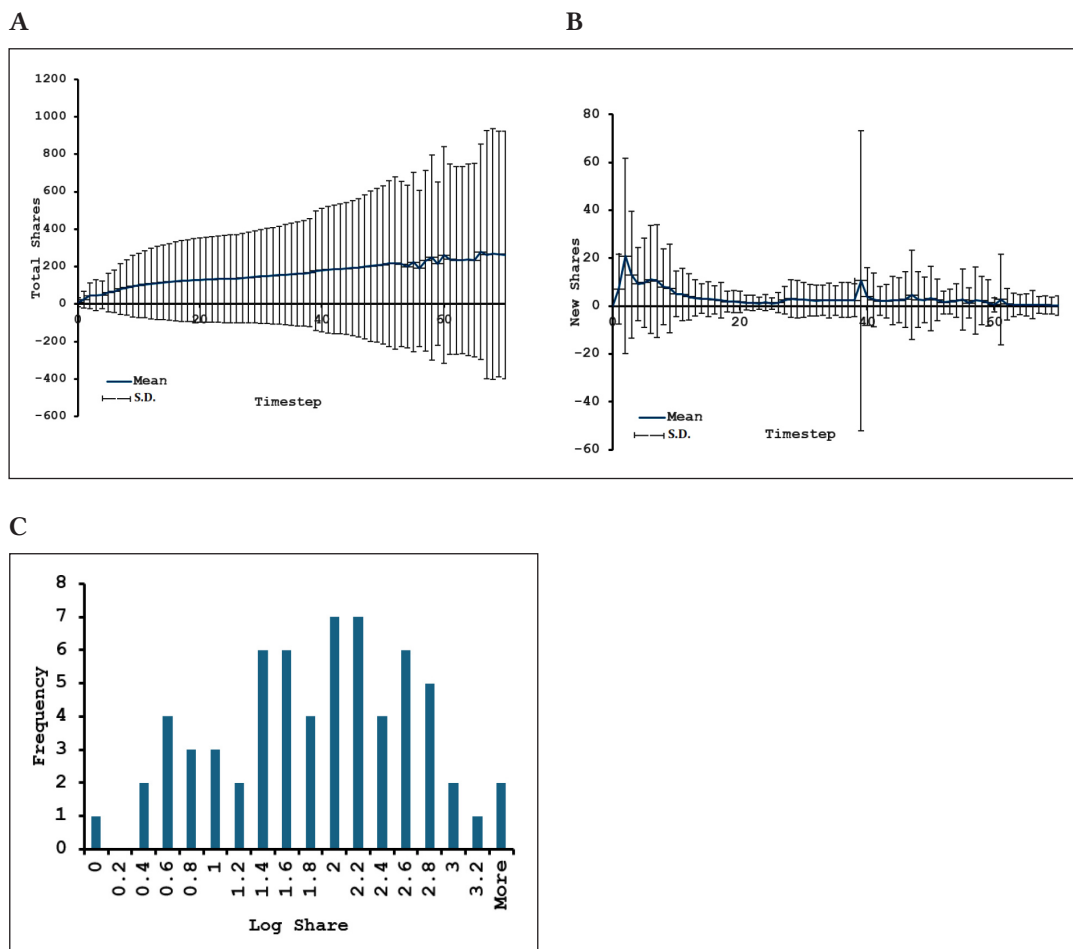
<sup>1</sup> <https://kozvelemenykutatok.hu/2020-januari-kutatasi-eredmenyek-idea/>



Crowdtangle data does not include any information about the private accounts that interact with certain posts; however, basic information is available about the Facebook page that created the post, such as the number of likes, number of followers, and country of posting. Thus, this information is available for every party and party leader's page. Crowdtangle data can be used in two ways: in the form of a summary of posts from a certain period from certain users or a detailed summary of all posts. In our analysis, we use both.

Throughout the period, 146 posts were collected from the 16 Facebook pages of the Hungarian party leaders. The collected data contains the number of reactions of each type ('like,' 'love,' 'care,' 'haha,' 'angry,' and 'sad') to posts from the parties and their leaders. The appendix contains the details of the posts published by these political actors.

### Statistical analysis



**Figure 1** A-B: Means and standard deviations of the total shares (A) and new shares (B) of posts in different time steps. C: Distribution of log(shares) after 70 time steps (16 days). Note: Crowdtangle time steps are not linear; steps 1-5 represent 15 minutes; steps 6-26, 30 minutes; steps 27-38, 1 hour; steps 39-46, 3 hours; steps 47-54, 6 hours; steps 55-60, 12 hours, and 61-70, 1 day.



The distribution of sharing behavior indicates that it is rather skewed; the median post was only shared 77 times. However, some successful posts were shared more than a thousand times. Therefore, we plotted the logarithm of total shares (Figure 1 C). Regarding sharing over time (Fig 1B), there is a rapid take-off period in sharing in the first two steps and then a gradual decrease. (The ‘bumps’ in Fig 1 A-B correspond to changes in the length of the time period represented in specific time steps, e.g., time represented by a time step doubles after step 5 and step 26).

After considering the distributions, we created a regression model to analyze sharing dynamics. Thus, New Shares for each post within a specific time step will be our dependent variable.

The New Shares variable is a discrete count variable for which research usually applies count regressions, most frequently Poisson or negative binomial regression. Negative binomial regression is a statistical method that is suitable for analyzing over-dispersed count data where the conditional variance exceeds the conditional mean. This condition is true for the New Shares variable, as Table 1 shows.

**Table 1** Description of New Shares variable

	Mean	SD
New Shares (N=9514)	3.3	13.7

As independent variables, we use Comments and Reactions over time to differentiate between overall user engagement and emotional reactions. In particular, we predict the quantity of New Shares with the quantity of New Reactions and New Comments in the previous Timestep. Additionally, the Time Period variable was added as an independent variable to the model to control for possible decreasing engagement over time, as we observe in Fig 1B. Furthermore, as we observe that engagement with specific posts is highly variable, we added post-specific fixed effects to the model to capture the differences in the attractiveness of the post. Correspondingly, we estimate the following regression

$$\log(\text{NewShare}_{i,t}) = \beta_0 + \beta_1 \text{NewReactions}_{i,t-1} + \beta_2 \text{NewComments}_{i,t-1} + \beta_3 t_i + \xi_i \quad (1)$$

Where  $t$  represents the specific Timestep,  $i$  is the indicator of the post, and  $\xi_i$  stands for the post-specific fixed effect.

Table 2 presents the results of the regression analysis corresponding to Equation 1. The findings indicate that the New Reactions variable has a significant positive effect on the number of New Shares, suggesting that reactions and shares are associated, even after controlling for the quality of the post according to fixed effects. Thus, positive feedback between these variables may be present in the social network. Comments, however, were not associated with such an effect. Time, furthermore, has a negative impact on new shares, even after controlling for the previous engagement. Because the possibility of multicollinearity of these variables arises, we tested for this. VIF values for the explanatory variables were in the moderate range (2.23, 2.25, and 1.02 for New Reactions, New Shares,

and Time, respectively). This regression framework corresponds to the decreasing engagement that we observed in Fig 1A-B. This phenomenon in the diffusion models comes from the saturation effect – the impact of the decreasing pool of agents susceptible to a new product or disease.

**Table 2** Regression analyses of New Shares by lagged Reactions, Comments, and Time Period

	<b>New Shares</b>
New Reactions (t-1)	0.006***
New Comments (t-1)	0.00
Time Period	-0.96***
Observations	9,329

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01.

### Agent-based model

In our agent-based model, we aim to analyze the diffusion of political news under different conditions with regard to network structure, homophily and echo-chamber effect, and the presence of a preference-based filtering algorithm. Building on diffusion models and our previous empirical analysis concerning the dynamics of sharing political news on Facebook, the simulation aims to create a simplified model of news sharing. Our observation is that reactions propagate future shares, and we incorporate them into the model by adding reactions as a separate channel, which increases the visibility of posts to friends. We consider the fixed effect term in the regression analysis to be the attractiveness of the post in the simulation. Corresponding to previous studies and because online social networks have been shown to exhibit properties predicted by different network models, three different network structures are simulated: random, small world, and preferential attachment. We consider these with or without homophily and with or without filtering algorithms to compare information-sharing under various network environments.

Considering the network structure, we have seen that the network structure of on-line social networks combines several features of basic models. Therefore, we follow the earlier literature and consider three network models for the simulation, being aware that none of these fully characterize real social networks. Under the random network condition, links between agents are formed probabilistically between nodes with uniform probability. Small world refers to a network that has high clustering (friends of friends tend to be friends) and relatively low average distances between nodes. High clustering is achieved by distributing the nodes on a circle and creating connections between each of them within a certain range on the circle. As the resulting network has high distances, in the next step, a small fraction of links are redistributed randomly to create ‘shortcuts.’ In preferential attachment networks, connections are distributed according to how many connections the node already has.

Echo chambers refer to the phenomenon of groups of like-minded users forming on social media and where there is a bias in the information diffusion toward like-minded users (Quattrociocchi et al., 2021). The homophily of users is one mechanism behind echo chambers, but social media algorithms may reinforce the effect (Cinus et al., 2022). Homophily refers to the tendency for connections to occur at a higher rate amongst those who share a common interest (McPherson et al., 2001).

In the simulation, homophily is operationalized as a higher number of links being simulated between those whose political interests are similar under the homophily condition. With respect to algorithms, a content-filtering algorithm was considered, whereby posts are shown with decreased probability to those users whose political attitude is more distant from the sender. To examine the potential interactions, each of the three types of network structures was simulated with and without homophily and with or without a preference-based filtering algorithm. The simulation was implemented using the Netlogo software package.

The news sharing in the simulation was implemented the following way. Each agent represents a person, a member of a social network. A fraction of agents are selected to be 'followers' of the politician; they are shown the information in the first timestep. Their neighbors are the connected nodes who can see their activities – reactions or shared posts. The number of neighbors of a given agent – the node degree – depends on the network structure. The sender (politician) is modeled as being external to the network. Political attributes are assumed to be one-dimensional: the politician stands at the zero point, and the agents are at different distances from it, modeled by a uniformly distributed 'interest' parameter. The politician posts different information having a random attractiveness parameter and political specificity. Sharing happens randomly based on the attractiveness of the post, its political specificity, and the distance between the agent and the politician on the political spectrum. (Specifically, the attractiveness parameter decreased by the political distance between the politician and the agent and by the political specificity of the post, which is evaluated against a random number). Reacting to the post happens similarly to sharing but with a higher probability.

Non-follower agents – agents that did not see the information in the first timestep – are only shown the post if their neighbor shared it or reacted to it. This, however, is not automatic. Posts shared (or reacted to) by friends are made visible to users randomly based on their political attitudes. Specifically, their distance from the politician decreased by the political specificity of the post is evaluated versus a threshold. We manipulate the content-filtering algorithm using this threshold. In the baseline case, almost everyone can see the post, while in the 'filtering algorithm' scenario, only those whose attitude is close to that of the politician can see it.

Thus, the simulation consists of the following steps:

1. A [Random / Preferential Attachment / Small World] network of people is created, having different attitudes ('interest') towards the sender [with / without] homophily. Some people are selected to be followers of the sender. The attractiveness of the post is defined.
2. Followers are set as eligible to see the post.

3. The post is shown to those who are eligible to see it and have attitudes close enough to the sender. In the filtering algorithm scenario, a strict threshold is applied; without the filtering algorithm, this threshold is loose.
3. Those who have seen the post and have not shared it yet decide if they will share it based on their attitude towards the post and the attractiveness of the post.
4. Those who have seen the post and have not reacted yet decide if they will react to it based on their attitude towards the post and the attractiveness of the post.
6. The connections of those who have shared the post or reacted to it are set as eligible to see the post.
7. The cycle starts over from Step 2.

In our baseline simulation, the network consists of 400 nodes, of whom 20 are followers of the politician, and 380 are not. Each node has an average degree of 4. Additionally, the simulation was repeated with different settings to test the robustness of the results. Table 3 summarizes the different settings for the simulation. The number of nodes and average node degree were modified, and in the case of small-world networks, the probability of the rewiring of the network was tested using two versions.

**Table 3** Settings for the agent-based simulations

Scenario	Network Type	Number of Nodes	Average Node Degree	Number of Followers
A	Small world ( $p = 0.1$ )	400	4	20
	Preferential Attachment	400	4	20
	Random	400	4	20
B	Small world ( $p = 0.1$ )	1000	4	50
	Preferential Attachment	1000	4	50
	Random	1000	4	50
C	Small world ( $p = 0.1$ )	1000	10	50
	Preferential Attachment	1000	10	50
	Random	1000	10	50
D	Small world ( $p = 0.05$ )	400	4	20

In Scenario B, we increased the network size from 400 to 1000 nodes, holding the average node degree constant. The number of followers was also increased proportionally to maintain the same ratio as before. In Scenario C, the number of nodes and the node degree were also increased by a factor of 2.5 compared to the original setup, resulting in a similar network density but a higher average degree. Last, Scenario D specifically targets small-world networks, testing a modification to the rewiring probability.

## 4 Results of the simulation model

### 4.1 Baseline specification

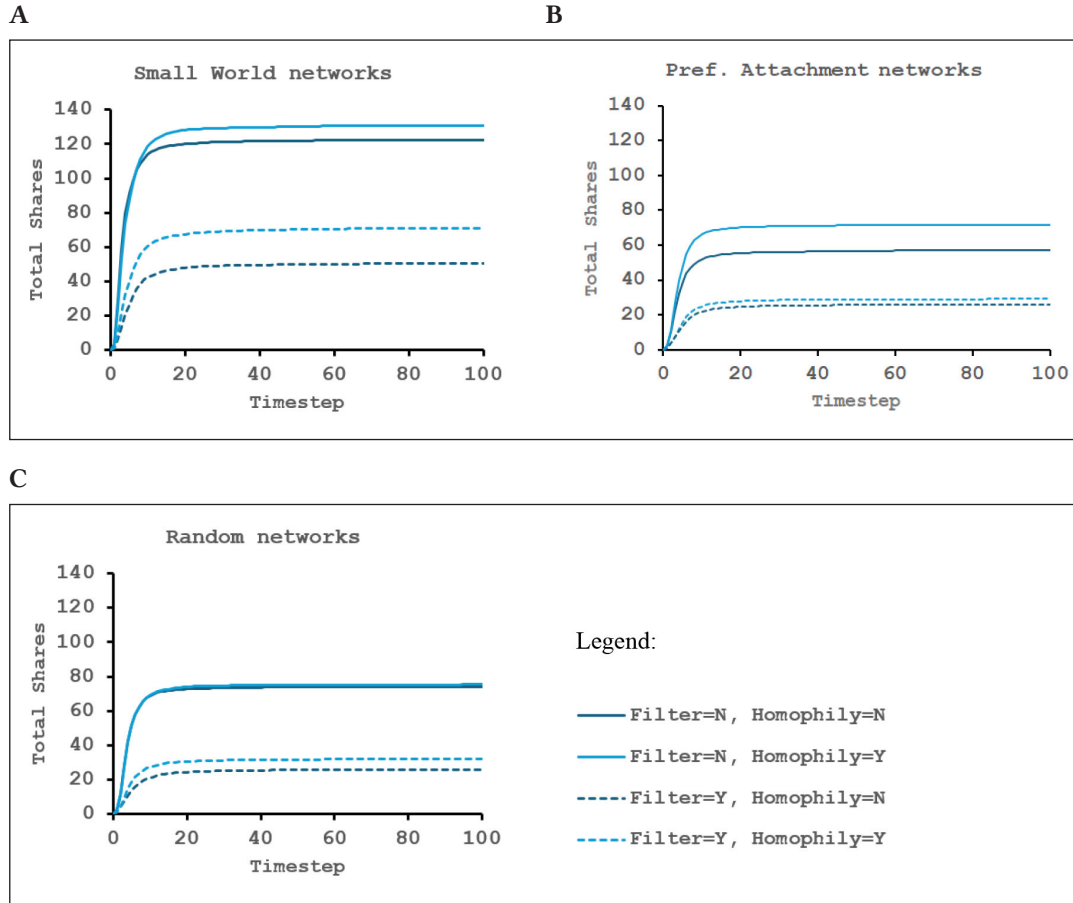
Table 4 presents the results of the baseline version of the ABM. The number of agents who have shared, watched, and reacted to the post in the 100th step of the simulation is presented. Results show that with each network type, a filtering algorithm decreases the number of agents who interact with the post and its final reach. Additionally, homophily tends to increase interaction with the post if a filtering algorithm is present. Results indicate that the highest average count of individuals who viewed, reacted to, or shared a post occurred within small-world networks, particularly those without filtering algorithms and incorporating homophily. On the contrary, the lowest count of nodes engaging in sharing was observed in random and preferential attachment networks, particularly those without homophily and with the presence of a filtering algorithm.

**Table 4** Average number of agents at each step who watched/reacted to/shared the post in the 100th step of 100-iteration Scenario A

Network Type	Homophily	Filtering algorithm	Number of nodes	Average node degree	Number of followers	Watched	Reacted	Shared
Small world	false	false	400	4	20	215.5	139.3	122.4
Small world	false	true	400	4	20	133.1	59.9	50.6
Small world	true	false	400	4	20	205.9	152.4	130.8
Small world	true	true	400	4	20	132.6	84.4	70.8
Random	false	false	400	4	20	180.2	103.5	74.2
Random	false	true	400	4	20	95.7	37.8	25.7
Random	true	false	400	4	20	160.5	105.9	75.3
Random	true	true	400	4	20	100.7	47.0	32.1
PA	false	false	400	4	20	156.4	77.4	57.0
PA	false	true	400	4	20	88.2	36.3	26.1
PA	true	false	400	4	20	163.0	99.4	71.7
PA	true	true	400	4	20	92.6	43.3	29.2

In Figure 2, we focus on sharing and visualizing the dynamics of Total Shares under these conditions. First, it is reassuring that the general shape of the diffusion curves is similar to what we observed in Figure 1A about the sharing of political content on Facebook. We can also visually observe what we have seen from Table 4 – that small-world networks

without filtering are the ones with the most efficient spreading, and homophily, in general, supports spreading. However, its effects are highly variable across the specifications.



**Figure 2** Average number of shares over the 100 steps of the simulations in different networks: A. Small World networks B. Preferential Attachment networks C. Random networks. -

To test the statistical significance of the influence of the different network attributes, a linear regression was applied to the number of nodes that shared the information in the simulation. The number of agents that shared the original information was explained by the network type, homophily, the presence of filtering algorithms, and the pairwise interactions of these factors. Accordingly, the following equation was estimated:

$$\begin{aligned}
 TotalShare = & b_0 + b_1 SmallWorld + b_2 PA + b_3 Random + b_4 Homophily + b_5 FilterAlgorithm + \\
 & b_6 Homophily \times FilterAlgorithm + b_7 PA \times Homophily + b_8 Random \times Homophily + \\
 & b_9 SmallWorld \times Homophily + b_{10} PA \times FilterAlgorithm + b_{11} SmallWorld \times FilterAlgorithm + \\
 & b_{12} Random \times FilterAlgorithm + \varepsilon
 \end{aligned} \tag{2}$$

Regarding network type, the coefficients of small world and preferential attachment (PA) networks were measured, and random networks served as a reference category, as the linear regression contained the network types as dummy independent variables.

**Table 5** Linear regression analysis of network attributes' influence on the total number of agents sharing posts in simulation Scenario A

	<b>B</b>
Intercept	181.82***
Small World	-16.81
Preferential Attachment	-54.21*
Homophily	-29.99
Echo chamber	-90.93***
Homophily and Echo chamber	26.77*
Homophily in Preferential Attachment	12.61
Homophily in Small World	19.38
Echo Chamber in Preferential Attachment	17.8
Echo Chamber in Small World	14.69
Adjusted R2	0.07943
Multiple R2	0.08634
Observations	1200

Note: \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$

Table 5 shows the estimation results corresponding to Equation 2. It shows that after one hundred timesteps, significantly fewer agents share the information in preferential attachment networks than in random networks, corresponding to the earlier descriptive finding. However, the regression analysis does not find that news spreads significantly more rapidly the small world network than in the random ones. Although both homophily and filtering algorithms may be viewed as mechanisms that restrict diffusion, homophily did not cause a statistically significant decrease in shares on its own. (In fact, in the descriptive statistics, it seemed to have a positive impact instead). The filtering algorithm was shown to have a negative impact on the regression analysis, similar to the descriptive results. Further, the regression model shows that the presence of homophily counteracts the negative effect of filtering algorithms; if both were introduced in the networks, this resulted in more total shares.



### Alternative specifications

Table 6 displays the result of the simulations run over a bigger network but with the same average degree. From the average statistics, we can observe that, in general, the news spreads more or less similarly in the three network types if there is no filtering algorithm. With filtering algorithms, however, small-world networks seem to perform somewhat better. Further, homophily supports spreading in cases when the filtering algorithm is present (in all network types), but when it is not present, its impact depends on the network type.

**Table 6** Average number of agents at each step who watched/reacted to/shared the post in the 100th step of 100-iteration Scenario B

Network Type	Homophily	Filtering algorithm	Number of nodes	Average node degree	Number of followers	Watched	Reacted	Shared
Small world	false	false	1000	4	50	558.0	295.2	268.5
Small world	false	true	1000	4	50	466.4	216.8	191.7
Small world	true	false	1000	4	50	555.4	407.0	365.5
Small world	true	true	1000	4	50	443.3	308.6	273.8
Random	false	false	1000	4	50	524.0	345.9	282.9
Random	false	true	1000	4	50	331.2	149.5	115.9
Random	true	false	1000	4	50	595.4	441.3	353.6
Random	true	true	1000	4	50	386.4	246.9	192.2
PA	false	false	1000	4	50	537.1	318.7	264.9
PA	false	true	1000	4	50	334.1	150.2	118.8
PA	true	false	1000	4	50	517.1	354.8	289.2
PA	true	true	1000	4	50	321.3	193.2	151.5

The results of the simulations corresponding to Scenario C, a network similarly large to Scenario B but with higher density, are presented in Table 7. Compared to the previous scenario, we see that the higher density somewhat increased the number of shares in the small-world network but not in the other two types. This results in small-world networks seeming to outperform the other two networks in each scenario. In this setting, the Total Shares after 100 steps were, on average, higher in each scenario that included homophily, compared to the similar setting without homophily.

**Table 7** Average number of agents at each step who watched/reacted to/shared the post in the 100th step of 100-iteration Scenario C

Network Type	Homophily	Filtering algorithm	Number of nodes	Average node degree	Number of followers	Watched	Reacted	Shared
Small world	false	false	1000	10	50	703.8	410.9	364.9
Small world	false	true	1000	10	50	472.6	232.2	200.2
Small world	true	false	1000	10	50	563.5	466.8	416.7
Small world	true	true	1000	10	50	501.7	393.1	344.4
Random	false	false	1000	10	50	612.3	358.3	265.1
Random	false	true	1000	10	50	470.6	199.3	134.7
Random	true	false	1000	10	50	478.8	331.9	232.5
Random	true	true	1000	10	50	392.6	285.0	199.6
PA	false	false	1000	10	50	549.3	299.7	224.4
PA	false	true	1000	10	50	415.0	180.8	126.6
PA	true	false	1000	10	50	562.0	386.1	286.2
PA	true	true	1000	10	50	362.5	237.0	164.8

Regarding the small world network, the simulations with the increased rewiring probability, that is, with more ‘distant’ ties but on the same network size as our baseline (Scenario A), are presented in Table 8. The number of shares at the end of the simulation somewhat increased in the new version compared to the original one (Table 4) when a new filtering algorithm is applied, but no change can be observed when filtering limits the visibility of posts. Similarly to the original scenario, homophily tends to mitigate the limiting effect of the filtering algorithm.

**Table 8** Average number of agents at each step who watched/reacted to/shared the post in the 100th step of 100-iteration Scenario D

Network Type	Homophily	Filtering algorithm	Number of nodes	Average node degree	Number of followers	Watched	Reacted	Shared
Small world	false	false	400	4	20	242.1	148.7	129.3
Small world	false	true	400	4	20	139.5	62.6	52.8
Small world	true	false	400	4	20	197.6	150.7	130.1
Small world	true	true	400	4	20	183.0	115.9	96.0

Comparing the different scenarios, the following tendencies can be observed. First, it is visible that the filtering algorithm limits the spread of news in the network, which is not surprising given that this mechanism directly limits the visibility of the post to users having different political preferences from the sender. Second, small-world networks seem to have an advantage in spreading the news. However, this is not consistent across different network structures and sizes in general. What seems consistent is that these networks perform better if there is a filtering algorithm. Third, homophily tends to act as a facilitator of diffusion in contrast to a limiting factor. This tendency is also not consistent across all network structures and sizes, but it is present in all specifications when the filtering algorithm is present.

After observing these tendencies, we should also check if the above differences are systematically present and statistically significant over the simulation runs. To do this, we re-run the linear regressions corresponding to Equation 2 on the alternative specifications.

The results of the regressions are summarized in Table 9. Columns B – D represent the corresponding alternative specification, while we included the results of the baseline specification again in column A for a clearer overview. The regression analysis only partly supports the descriptive tendencies described above. The tendency that, in the case of filtering algorithms, small world networks are more efficient in terms of spreading the news (as compared to random networks) is supported in scenarios A and B but not in the bigger, denser network (Scenario C). The tendency for homophily to counteract the negative impact of the echo chamber is supported in the original scenario and also in the small world network with more distant ties (Scenario D), but not in the bigger network (Scenario B), while in the case of Scenario C, the positive effect of homophily is specific to small-world networks.

**Table 9** Linear regression analysis of the influence of network attributes on the total number of agents sharing posts in Scenarios A, B, C, and D

Scenario	A	B	C	D
Intercept	181.82*** (33.72)	371.5*** (103.35)	430.85*** (101.76)	247.75*** (61.36)
Small World	-16.81 (30.65)	-105.2075 (93.93)	56.28 (92.48)	
Preferential Attachment	-54.21* (30.65)	4.07 (93.93)	-16.14 (92.48)	
Homophily	-29.99 (19.89)	73.995 (60.95)	-66.94 (60.01)	-41.82 (38.81)
Echo chamber	-90.93*** (19.89)	-163.725** (60.95)	-164.81** (60.01)	-119.14** (38.81)
Homophily and Echo chamber	26.77* (11.48)	-0.3233 (35.19)	55.44 (34.65)	42.55* (24.54)
Homophily in Preferential Attachment	12.61 (14.06)	-45.06 (43.09)	33.76 (42.44)	
Homophily in Small World	19.38 (14.06)	16.035 (43.09)	81.75* (42.44)	
Echo Chamber in Preferential Attachment	17.8 (14.06)	22.31 (43.09)	-27.98 (42.44)	
Echo Chamber in Small World	14.69 (14.06)	79.925* (43.09)	-36.89 (42.44)	
Multiple R2	0.07943	0.05882	0.0711	0.05571
Observations	1200	1200	1200	400

Note: \*p < 0.05; \*\*p < 0.01; \*\*\*p < 0.001

## 5 Discussion

Online politics influence offline political action through social media. Research has shown that offline mobilization in online spaces is most effective through the sharing of political content with people we know. In this article, we sought to explore the logic of news sharing on social media.

Our aim was to examine the impact of factors that previous research has shown to influence sharing: network structure, homophily and echo chambers, and filtering algorithms. To base our agent-based model, we analyzed Facebook posts of Hungarian politicians from 2021. However, beyond the fact that Facebook dominated the social media landscape, the field of study is not specific. In the social media environment, information

diffusion is influenced by an algorithm that is linked to friends' interactions with the information (Bucher, 2012). Interactions regarding a post on Facebook can involve expressing emotions via reaction buttons, commenting, and sharing. The focus of the analysis is sharing. This was explained by other reactions in the regression model, and it showed a significant effect on the number of shares. Thus, it was implanted into an agent-based model.

Because online social networks share various characteristics of network models, corresponding to the literature, three network models – small world, preferential attachment, and random networks – were compared. The simulation confirmed the importance of weak ties in social networks; most people were reached by shared news in small-world network scenarios, corresponding to the findings of prior research (Centola, 2010; Pegoretti et al., 2012). A novel element in the analysis is, however, that small-world networks overperformed preferential attachment (Albert & Barabási, 2002) networks, too, in terms of the final reach of information. When comparing the different network scenarios, it was observed when network density was relatively low and/or the spread of the news was constrained by a filtering algorithm.

This result may be due to two features of the model. First, sharing is not automatic; it depends on the political alignment of users, who are heterogeneous in this regard. Second, there is no social influence in our model. Therefore, if a central person is very negative towards a politician, they will not share the information, even if many of their friends do. Thus, high centralization in our model may stop the diffusion process if the central person happens to be skeptical, while in a less centralized network, like in a random or small-world network or the case of denser preferential attachment networks, the information bypasses the skeptical person more easily.

The feature of the heterogeneous thresholds of the agents is typical in diffusion models. Otherwise, the diffusion question would be reduced to the question of the shortest path in the network. The lack of social influence is not that typical; such a mechanism is included in several related models. In opinion dynamics models, e.g., Baumann et al. (2020), the basic assumption is that neighbors in networks influence each other. In network models of collective action, like a protest, for example (Chwe, 1999; Korkmaz et al., 2019), individuals favor acting only if enough others act similarly. This setup is similar to the adoption of innovations if the innovation includes a network externality because its utility comes from connecting people, like a messaging system (Aral et al., 2013). The lack of a social influence mechanism in our model also means that the advantage of small-world networks associated with locally dense structures (high transitivity) that ensure collective action and network-based diffusion (Pegoretti et al., 2012) is not present in our model.

The emergence of echo chambers has been at the forefront of social media research recently (Del Vicario et al., 2016; Quattrociocchi et al., 2021), which this article approaches by introducing homophily to the networks together with a bias in the social media algorithm that filters content according to its fit with the attitudes of the agents. In general, we find that the impact of homophily varies across networks depending on the structure and specification, similar to Korkmaz et al. (2019). An interesting result of our simulation is that in the cases when the diffusion of news is limited by a content-filtering algorithm, homophily enhanced diffusion of the news, especially in small-world networks, instead of

limiting it. This result contrasts with what we had expected based on earlier studies arguing that homophily and computer algorithms amplify each other in creating echo chambers (B. Jiang et al., 2021) and showing that diverse connections boost diffusion (Cota et al., 2019). About this result, it can be argued that in the case of bounded diffusion opportunities, when the news itself is not very attractive, and the filtering algorithm does not allow it to be seen by politically distant agents, homophily does not act as a limiting factor, but as an enhancing factor of diffusion. This happens because the connections between similar people create a path through which the politically interested agents can be reached by the news. In this setting, without homophily, the spread of the news stops early, and agents who are politically interested but distant in the network from the source are not reached. Results show that homophily has a positive effect when echo chambers are present, especially in network types where diffusion is low on average, supporting this interpretation. Thus, in other words, echo chambers have the function of enhancing diffusion in the case of news that has a more limited potential audience. Note that we find this without taking into account the attention constraints of users, under which condition filter bubbles have the additional function of creating a stream of potentially attractive content that users are willing to share.

## References

- Aral, S., Muchnik, L., & Sundararajan, A. (2013). Engineering social contagions: Optimal network seeding in the presence of homophily. *Network Science*, 1(2), 125–153.
- Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020). Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, 124(4), 048301.
- Bene, M. (2018). Post shared, vote shared: Investigating the link between Facebook performance and electoral success during the Hungarian general election campaign of 2014. *Journalism & Mass Communication Quarterly*, 95(2), 363–380.
- Bene, M., & Somodi, D. (2018). “Mintha lenne saját médiánk...” A kis pártok és a közösségi média. *Médiakutató*, 19(2), 7–20.
- Boulianne, S. (2020). Twenty Years of Digital Media Effects on Civic and Political Participation. *Communication Research*, 47(7), 947–966. <https://doi.org/10.1177/0093650218808186>
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science*, 329(5996), 1194–1197.
- Chwe, M. S. (1999). Structure and Strategy in Collective Action. *American Journal of Sociology*, 105(1), 128–156. <https://doi.org/10.1086/210269>
- Cinus, F., Minici, M., Monti, C., & Bonchi, F. (2022). The effect of people recommenders on echo chambers and polarization. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 90–101.
- Cota, W., Ferreira, S. C., Pastor-Satorras, R., & Starnini, M. (2019). Quantifying echo chamber effects in information spreading over political communication networks. *EPJ Data Science*, 8(1), 35.

- Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Echo chambers: Emotional contagion and group polarization on Facebook. *Scientific Reports*, 6(1), 37825.
- Dimitrova, D. V., Shehata, A., Strömbäck, J., & Nord, L. W. (2014). The effects of digital media on political knowledge and participation in election campaigns: Evidence from panel data. *Communication Research*, 41(1), 95–118.
- Feezell, J. T., Conroy, M., & Guerrero, M. (2016). Internet use and political participation: Engaging citizenship norms through online activities. *Journal of Information Technology & Politics*, 13(2), 95–107.
- Ge, Y., Zhao, S., Zhou, H., Pei, C., Sun, F., Ou, W., & Zhang, Y. (2020). Understanding echo chambers in e-commerce recommender systems. *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2261–2270.
- Groshek, J., & Dimitrova, D. (2011). A cross-section of voter learning, campaign interest and intention to vote in the 2008 American election: Did Web 2.0 matter. *Communication Studies Journal*, 9(1), 355–375.
- Helbing, D. (2012). *Agent-based modeling*. Springer.
- Holt, K., Shehata, A., Strömbäck, J., & Ljungberg, E. (2013). Age and the effects of news media attention and social media use on political interest and participation: Do social media function as leveller? *European Journal of Communication*, 28(1), 19–34.
- Jiang, B., Karami, M., Cheng, L., Black, T., & Liu, H. (2021). Mechanisms and Attributes of Echo Chambers in Social Media. *arXiv Preprint arXiv:2106.05401*.
- Jiang, Y., & Jiang, J. C. (2014). Diffusion in social networks: A multiagent perspective. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(2), 198–213.
- Klinger, U. (2013). Mastering the Art of Social Media. *Information, Communication & Society*, 16(5), 717–736. <https://doi.org/10.1080/1369118X.2013.782329>
- Koltai, J., & Stefkovics, Á. (2018). A big data lehetséges szerepe a pártpreferenciabecslésekben magyarországi pártok és politikusok Facebook-oldalainak adatai alapján. Módszertani kísérlet. *Politikatudományi Szemle*, 27(2.), 84–120.
- Korkmaz, G., Kuhlman, C. J., Goldstein, J., & Vega-Redondo, F. (2019). A computational study of homophily and diffusion of common knowledge on social networks based on a model of Facebook. *Social Network Analysis and Mining*, 10(1), 5. <https://doi.org/10.1007/s13278-019-0615-5>
- Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and Control in Modern Society*, 18(1), 18–66.
- Li, Z., & Tang, X. (2015). From global polarization to local social mechanisms: A study based on ABM and empirical data analysis. *Smart Modeling and Simulation for Complex Systems: Practice and Theory*, 29–40.
- MacWilliams, M. C. (2015). Forecasting congressional elections using Facebook data. *PS: Political Science & Politics*, 48(4), 579–583.



- Magin, M., Podschuweit, N., Haßler, J., & Russmann, U. (2017). Campaigning in the fourth age of political communication. A multi-method study on the use of Facebook by German and Austrian parties in the 2013 national election campaigns. *Information, Communication & Society*, 20(11), 1698–1719.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.
- Oser, J., Hooghe, M., & Marien, S. (2013). Is online participation distinct from offline participation? A latent class analysis of participation types and their stratification. *Political Research Quarterly*, 66(1), 91–101.
- Pegoretti, G., Rentocchini, F., & Marzetti, G. V. (2012). An agent-based model of innovation diffusion: Network structure and coexistence under different information regimes. *Journal of Economic Interaction and Coordination*, 7(2), 145–165.
- Quattrocioni, W., Cinelli, M., De Francisci Morales, G., Galeazzi, A., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9).
- Skoric, M. M., & Zhu, Q. (2016). Social media and offline political participation: Uncovering the paths from digital to physical. *International Journal of Public Opinion Research*, 28(3), 415–427.
- Theocharis, Y., & Lowe, W. (2016). Does Facebook increase political participation? Evidence from a field experiment. *Information, Communication & Society*, 19(10), 1465–1486.
- Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.
- Ylikoski, P. (2014). Agent-based simulation and sociological understanding. *Perspectives on Science*, 22(3), 318–335.
- Zhang, W., Johnson, T. J., Seltzer, T., & Bichard, S. L. (2010). The revolution will be networked: The influence of social networking sites on political attitudes and behavior. *Social Science Computer Review*, 28(1), 75–92.

## Appendix

Name	Number of posts	Number of Followers at Posting	Number of Likes at Posting	Total Likes	Total Comments	Total Shares	Total Love	Total Wow	Total Haha	Total Sad	Total Angry	Total Care	Total Interactions
Demokratikus Koalíció	8	1317882	1177101	9210	581	1645	132	52	57	250	166	108	12201
Dr. Tóth Bertalan	2	70804	67440	8222	599	4428	50	116	213	697	990	53	15368
Fekete-Győr András	5	200855	186283	4413	2219	628	85	43	544	100	299	29	8360
Fidesz	28	8923704	8867640	38805	1465	5200	273	46	225	168	341	377	46900
Gyurecsány Ferenc	2	578164	534622	3130	1191	839	343	6	111	15	42	148	5825
Jakab Péter	5	1182516	944805	13526	1692	1862	164	35	129	510	1033	145	19096
Jobbik Magyarországért Mozgalom	15	7798134	7729951	12335	3913	7807	75	389	4250	950	5431	49	35199
Kanász-Nagy Máté	6	31329	30612	227	35	65	1	5	7	7	4	3	354
Karácsony Gergely	7	1892085	1822456	39745	4112	4357	2078	584	995	105	287	1227	53490
Kunhalmi Ágnes	2	84397	82318	1164	252	30	77	0	4	0	0	169	1696
LMP – Magyarország Zöld Pártja	10	740252	763466	209	38	69	4	5	10	5	24	1	365
Magyar Szocialista Párt	17	3668436	3541908	9937	694	3342	125	28	50	179	400	153	14908
Momentum Mozgalom	14	1621291	1512826	3894	437	301	143	28	163	30	147	66	5209
Orbán Viktor	4	4435914	3620652	32417	5178	741	632	57	1324	20	295	439	41103
Párbeszéd Magyarországért	14	1835916	1776396	7224	1363	1761	22	74	446	1768	2030	63	14751
Schmuck Erzsébet	7	223583	242210	311	27	35	1	3	2	3	13	2	397
Szabó Tímea	0	-	-	0	0	0	0	0	0	0	0	0	0
Total	146	34605262	32900686	184769	23796	33110	4205	1471	8530	4807	11502	3032	275222

**Appendix 1** Summary of Facebook posts collected (Source: Crowdtangle)

---

RADU M. MEZA\* & ANDREEA MOGOS\*\*

## Fear and Loathing in Eastern Europe. Framing and Agenda Setting of Foreign Affairs in Czech, Polish, and Romanian Sputnik News (2017–2022) – An Analysis of Headlines

Intersections. EEJSP

11(1): 63–85.

<https://doi.org/10.17356/ieejsp.v11i1.1229>

<https://intersections.tk.hu>

---

\* [radu.meza@ubbcluj.ro](mailto:radu.meza@ubbcluj.ro) (Babes-Bolyai University)

\*\* [andreea.mogos@ubbcluj.ro](mailto:andreea.mogos@ubbcluj.ro) (Babes-Bolyai University)

### Abstract

The Sputnik News network of websites has been the subject of scholarly attention since its 2014 launch in the context of the beginning of the Russo-Ukrainian conflict. In March 2022, Sputnik News websites were blocked by EU countries, thus acknowledging the network's potential influence. Some researchers claim that Sputnik news follows the model of Soviet propaganda, while others claim its discursive patterns point towards new communication strategies. There is scarce literature on Eastern European Sputnik editions and no comparative approaches to date. This research article aims to fill that gap by employing mixed methods to study a large dataset (N=118,198) of Sputnik News headlines from the foreign news sections of three language editions – Czech, Polish, and Romanian, from January 2017 to January 2022. The main findings indicate a conflict-oriented, war journalism approach that mostly focuses on USA-Russia and the individual actors Donald Trump and Vladimir Putin. Furthermore, the coverage leans towards fear as the dominant affective framing, and the most frequently represented organization, NATO, is framed as a failing alliance, actively pursuing the consolidation of power or even as an aggressor.

**Keywords:** Sputnik News, affective framing, agenda setting, war journalism, computational analysis, mixed methods.

## 1 Introduction

In the aftermath of the Russian Federation's annexation of Crimea, the Rossiya Segodnya state agency reformed Voice of Russia radio broadcasting and RIA Novosti news agency into Sputnik – a Russian-government-owned multiplatform news agency, multilingual news website network and radio broadcast service. Sputnik had worldwide editions in English and Spanish, European editions in German, French, Greek, Italian, Czech, Polish, Serbian, Latvian, Lithuanian, 'Moldavian' (sic), and Belarusian and Transcaucasian (Armenian, Abkhaz, Ossetian, Georgian, Azerbaijani) and Middle Eastern (Arabic, Turkish, Persian). In March 2022, Sputnik content was banned across the European Union.

---

Since Sputnik first appeared in the context of the onset of the Crimea crisis, it is worthwhile investigating how official Russian geopolitical narratives framed the rising tensions in the region. This research analyzes the three editions published in the languages of the countries directly bordering Ukraine – Poland (Polish), Slovakia (Czech), Romania, and Moldova (Romanian).<sup>1</sup>

The news reporting of Sputnik News and Russia Today (RT) often faced scrutiny due to ownership by the Russian state. Their funding has attracted international attention, with sources claiming annual budget allocations of over 300 million euros for RT and over 100 million euros for Rossiya Segodnya, which runs Sputnik News (EUvsDisinfo, 2019; US Department of State Global Engagement Center, 2022). RT can be considered a foreign policy tool of the Russian government, and it has played a significant role in spreading conspiracy theories in past years (Yablokov, 2015). Sputnik News' role manifests in shaping the news, constructing strategic narratives to advance Russia's public diplomacy efforts, and using crisis situations as a geopolitical instrument (Deverell et al., 2021; Demjanski, 2020; Mogoş et al., 2022). As compared to RT, which gained notoriety as an instrument of Russian propaganda, Sputnik News is considered to use gentler methods of persuasion (Furman et al., 2023). According to Stefan (2020), Sputnik News is a key channel used by the Kremlin to spread disinformation campaigns beyond Russia's borders, impacting the European Union, its Member States, and neighboring countries until the beginning of the war in Ukraine on February 24, 2022, when EU banned Sputnik websites and RT operations in Europe to address concerns about disinformation about Ukraine.

Recent geopolitical events and the global health crisis of 2020–2021 have sparked interest in researching the narratives put forward by sources such as Sputnik News. The representation of the COVID-19 pandemic and vaccine technologies have been the focus of research (Furman et al., 2023; Mogoş et al., 2022) associated with the Turkish and Romanian language editions, respectively. Research on Sputnik News coverage was also included in comparative framing and news slant research of specific topics such as the Israel and Palestine conflict (Shahzad, 2023). Other research looks at intermedia agenda setting between the English Sputnik News and US news sources and blogs (Wilbur, 2021) and finds that the criticism employed is often subtle. Watanabe (2017) analyzed conspiracy frames in the English-language Sputnik articles and found conspiracy frames, particularly in stories about the US and the UK, as primary targets of the respective language edition. Analysis of the general discursive patterns in Sputnik news editions other than English is scarce. In an analysis of the Swedish-language Sputnik News, Kragh and Åsberg (2017) found that the EU, NATO, and the US are the most frequent targets of criticism, with the EU 'depicted

<sup>1</sup> The 'Moldovan' / 'Moldavian' glossonym has been a subject of political debate since the break-up of the USSR. The 1991 Declaration of Independence of Moldova refers to its national language as Romanian, while the 1994 Constitution declared 'Moldovan' / 'Moldavian' as the official language of the Republic of Moldova. Later, a Constitutional Court ruling in 2013 interpreted the Declaration of Independence as superseding the Constitutional provision, thus acknowledging the official language glossonym as 'Romanian', with only the region of Transnistria still using 'Moldovan' as the name of the official language. The 'Moldovan' Sputnik edition hosts a considerable amount of news content targeting Romanian audiences (ro.sputnik.md), but there is no separate 'Romanian' edition. Czech and Slovak languages form a dialect continuum and are generally considered mutually intelligible. There is no Hungarian edition.

as an organization in terminal decline, beset by major crises' and NATO as 'a US instrument of war and the chief architect of Western policy towards Russia.' The analysis of the Swedish-language Sputnik identifies metanarratives such as (1) the failures and setbacks of Western policy, (2) negative representations of countries in the Western sphere of influence, (3) international cooperation and business ties with Russia, (4) Russian successes and a positive image of Russia, (5) dividedness between NATO allies, (6) Western/NATO aggressiveness/militarism, and (7) the West in crisis, aside from human interest stories or other news.

Although Eastern European countries of the former Eastern bloc are targeted through the respective language editions of Sputnik News, there has been no attempt at a comparative or integrated analysis of Sputnik News representations in Polish, Czech, and Romanian languages. In order to fill this gap in the literature, this research aims to explore the discursive patterns of coverage through the textual analysis of headlines published in the international news section of the Polish, Czech, and Romanian editions over five years (2017 to 2022). Poland, Czechia, and Romania have all been NATO members since the post-Cold War enlargement. Furthermore, since Sputnik News was created in the context of the conflict in Ukraine, the three countries are the closest NATO member countries in the region for which a Sputnik language edition exists.

## 2 Literature review

Officially a source of Russian public diplomacy, Sputnik has been used in previous research to provide 'a vision of the dominant Russian metanarratives' (Kragh & Åsberg, 2017) and a subsequent content analysis of the Swedish edition of Sputnik (removed from the Web in 2016 alongside all Nordic language editions) shows several consistent narratives that provide evidence of a blurring distinction between public diplomacy and what some authors call 'active measures' (a Soviet-era designator) – 'framing NATO as an aggressor and military threat, the EU as in terminal decline, and Russia as under siege from hostile Western governments' (Kragh & Åsberg, 2017). In recent scholarship, Sputnik has also been classified as 'alternative media with an affinity to populism' (Müller & Schulz, 2021). The populist angle may be explored in relation to the coverage of international leaders such as Donald Trump in the US or pro-Brexit campaign leaders and Viktor Orbán (Hungary) in Europe. Literature of the past decade has hardly been in agreement about Russia's propaganda and to what extent the associated messages and techniques differ from the Soviet era – with some authors (Jaitner & Mattson, 2015; Wilson, 2015; Yablokov, 2015) finding novelty in either the Russian approach or the contexts created by online media, while others arguing that the Soviet era dismiss-distort-distract-dismay approach to misrepresenting events (Snegovaya, 2015) still holds power. Based on a large-scale analysis of English-language Sputnik news, Watanabe (2018) claims that, in the aftermath of the 2014 Ukraine crisis, Sputnik has employed techniques different from those of Soviet-era propaganda: 'Russia is aiming to promote anti-establishment sentiment by publishing tailored conspiracy theories on Sputnik News websites' (Watanabe, 2018).

Recent literature on Russian propaganda proposes the concept of 'propaganda on demand,' defined as 'a kind of cynical political marketing in which narratives that trigger

or comfort certain social groups are purposefully used to manipulate public opinion' (Litvinenko, 2022). This concept, derived from the digital 'on-demand culture,' is characterized by the Russian government's use of digital media in conjunction with traditional media (Litvinenko & Toepfl, 2019), resulting in a communication strategy characterized by '(1) the multichannel distribution of propagandistic content, (2) large volumes of information or flooding, (3) an eclectic set of messages' (Litvinenko, 2022). Toepfl (2020) and Litvinenko (2022) argue that the proliferation of different, sometimes contradictory narratives that co-exist on digital platforms creates opportunities for members of the networked publics to select the narratives that best suit them. As opposed to coherent propaganda narratives, the emergent 'propaganda on demand' strategy is fluid and inconsistent, flooding multiple digital platforms with eclectic messages that are hard to counter and encouraging 'post-truth' stances that support the claim that 'everything is not so clear-cut' (Litvinenko, 2022). We examine the representation of foreign events by Sputnik News to Eastern European audiences in an attempt to identify the features of propaganda directed at foreign audiences.

This exploratory research is grounded in agenda-setting and framing theories. From the initial formulation of agenda-setting theory (McCombs & Shaw, 1972) to the model put forward in the Network Agenda-Setting Framework (Guo & McCombs, 2011), this media effects theory proposes that news media coverage patterns may transfer the salience of objects, attributes, and relationships among them from the media agenda to the public. Frames in a media text refer to the deliberate selection of certain aspects of reality, emphasized to convey a dominant interpretation. Media frames are employed to define issues, diagnose causes, induce moral judgments, or propose solutions for the issue being addressed (Entman, 1993). Analyzing news media frames through content analysis can be pursued using one of two distinct approaches: the inductive method, which involves an open exploration of dominant meanings found in media texts, and the deductive approach, which relies on pre-established variables (Semetko & Valkenburg, 2000). The past decades of research in communication science have seen the convergence of agenda-setting and framing models (McCombs & Ghanem, 2001), considering the conceptualization of news media frames as being 'manifested by the presence or absence of certain key-words, stock phrases, stereotyped images, sources of information, and sentences that provide thematically reinforcing clusters of facts or judgments' (Entman, 1993) that allow for text-based analytical approaches such as content analysis and discourse analysis, but also the network analysis of message content (Danowski, 1993) to be employed on news content.

We thus formulate the first research question: *RQ1: What are the most prominent actors (nations and public figures) and relations represented in Sputnik News headlines?*

According to Teun van Dijk (1993), (news) stories can be looked at as expressions of situation models, embodying the interpretation of an event and organized by a schema that features categories such as (1) Setting, (2) Participants, and (3) Actions. Narrative analysis applied in literary criticism and structuralist discourse analysis is of course much more complex (Herman & Vervaeck, 2019), but this research only endeavors to identify prominent actors and their attributed roles. Robertson (2017) proposes a coding scheme that allows distinctions to be made among narrative participants: (1) 'real' people; (2) abstractions like states or collective actors; (3) character types. Furthermore, micro-level narratives may be related to accumulated macro-level recurrent narrative themes: (1) 'public



narratives'—such as 'the working-class hero' or 'the enemy within'—or (2) 'master narratives'—'Progress', 'the triumph of Democracy' (Robertson, 2017). ). It is this type of macro-level narratives that Jack Lule (2001) puts forward through 'the seven master myths in the news': (1) 'the victim'; (2) 'the scapegoat'; (3) 'the hero'; (4) 'the good mother'/benefactor; (5) 'the trickster'; (6) 'the other world'; (7) 'the flood'. Similarly, Hayden White's four types of emplotment (White, 1973) inspired by Northrop Frye's (1957) theory of myths (mythoi) are also recognizable macro-structures that can be used to characterize the discursive patterns employed by historians: (1) comedy; (2) romance; (3) tragedy; and (4) satire/irony. Such categories have been applied in recent research in international relations (Kuusisto, 2018; 2019) as 'basic plots,' using the features associated with each of the four myths to classify stories about world politics. Using a discourse-mythological approach, Kelsey et al. (2023) employed the archetypal roles of mature masculinity proposed by Moore and Gillette (1991) to theorize archetypal blending in the modern representation of the army hero – from 'magician,' 'warrior,' and 'king' to 'the "global" hero-magician,' 'the "national" hero-warrior,' and 'the "domestic" hero-king.' We employ the concept of archetypal blending to analyze the framing of collective global actors such as NATO by Sputnik News. In the original conceptualization of roles, the main roles are each characterized by an apex and also by an active and a passive shadow role: the King, the Tyrant, and the Weakling; the Warrior, the Sadist, and the Masochist; the Magician, the Detached Manipulator, and the Denying 'Innocent' One. Recent research on Russian and Chinese media representations of NATO found two pairs of opposing frames: namely, NATO as an (1. a) aggressive or (1. b) defensive organization and as a (2. a) divided thus weak or (2. b) united thus strong organization. We formulate an additional research question to examine the most prominent actor that is identified: *RQ1a: What are the dominant frames used to represent NATO, the most prominent actor identified in the sampled editions of Sputnik News?*

Bednarek and Caple (Bednarek & Caple, 2014; Caple & Bednarek, 2016) look at linguistic indicators of news values that emphasize specific aspects of newsworthiness and function as framing devices. In the context of technological developments in the area of machine learning, methodological opportunities for using linguistic indicators in both inductive and deductive frame analyses have recently opened up, such as the computational framing analysis approach proposed by Guo et al. (2023). The 'war and peace' journalism framework was used by Maslog and Lee (2005) and subsequently (Nelson, 2019) to analyze the framing of conflict. Our research endeavors to detect linguistic indicators of the generic news frames proposed by Semetko and Valkenburg (2000) and the indicators of war and peace journalism employed by Nelson (2019) and Maslog and Lee (2005). Our second research question is thus: *RQ2: What textual indicators of generic frames and conflict coverage are employed?*

According to Wahl-Jorgensen (2020), a growing body of research in journalism studies is engaging with the concept of emotion, signaling 'an emotional turn.' Papacharissi (2015) introduced the concept of 'affective publics' to describe the context created by digital platforms, where the use of various digital means of aggregation and shared emotional expressions, opinions, and subjective experiences facilitate the construction of social movements and mobilization. News coverage is increasingly shaped by engagement with emotion, and '[a]udiences are more likely to be emotionally engaged, recall information and take action when news stories are relatable' (Wahl-Jorgensen, 2020). From the



researcher's perspective, the digital media landscape allows for data-driven sentiment analysis and emotion detection methods, the use of which rely on the assumption that 'emotions are inseparable from opinion, evaluation and decision-making' (Wahl-Jorgensen, 2020). Many news stories are simply re-packaged public statements or social media posts. News sources often rely on the colorful, emotional language selected from significant quotes by individual political actors to headline and elicit emotional engagement from the audience without any accountability for including value judgments.

The last research question is thus formulated as follows: *RQ3: What types of affective framing are employed in the Sputnik News representation of global events for Eastern European audiences?*

### 3 Methods

The research employs a large dataset (N=118,198) of Sputnik news headlines from three editions of Sputnik news targeted at audiences in EU, NATO member countries neighboring Ukraine: Czech/Slovak (NCZ= 46873), Polish (NPL= 54129), and Romanian/Moldavian (NRO= 16136). The data (news headlines, URL, and publishing date) were collected from the international news sections of each website: Czech (CZ: <https://cz.sputniknews.com/svet/>), Polish (PL: <https://pl.sputniknews.com/swiat/>), and respectively Romanian (RO: <https://ro.sputnik.md/International/>) for a period of five years – from January 2017 to January 2022, using Helium Scraper software.

The study focuses on headlines after a thorough examination of the structure and characteristics of Sputnik articles, which in most cases propose narrative headlines that reveal the actor/target, the action taken and/or the effect/outcome, in many cases accompanied by the moral judgement – these elements being sufficient for a reliable coding process.

For some parts of the analysis, headlines were translated into English using automated translation tools such as Google Translate API and Reverso ([reverso.net](https://reverso.net)). Pinpoint (<https://journaliststudio.google.com/pinpoint/>) was used to identify named entities – persons, organizations, and locations – as a preliminary step in developing lexicons for each of the languages, in order to perform the co-occurrence analysis.

This mixed methods research combines the use of machine-based learning tools with quantitative content analysis tools and qualitative analysis approaches on a limited subsample of the dataset.

#### 3.1 Computational Approach & Quantitative Content Analysis

The data was analyzed using KH Coder (Higuchi, 2016), Orange (Demšar, 2013) with an add-on for emotion and mood state categorization of short text trained on English-language tweets (Colnerič & Demšar, 2016) and WordStat 9. The emotion categorization model that was used was the six basic emotions associated with facial expressions identified by Eckman (Ekman & Friesen, 1971; Ekman, 1992).

To answer the first research question, codes were defined in each of the respective languages based on Named Entity Recognition (NER) provided by Google Pinpoint. Coding was automated through KH Coder and co-occurrence analysis was performed on the

three distinct language subsets. For the Czech and Polish datasets, a threshold of a minimum of 50 occurrences was used as a relevance criterion, while for the Romanian the threshold was set at 20 occurrences, proportionate to the smaller dataset size.

For the second research question, a translated version of the full dataset was used, and categories were defined deductively and inductively based on frequent words and phrases (bigrams and trigrams) occurring in more than 50 headlines from the dataset (N=118,198). Terms and key phrases were categorized into (1) the five generic frames defined by Semetko and Valkenburg (2000), (2. a) war journalism framing, and (2. b) peace journalism framing with a set of categories constructed based on previous literature (Nelson, 2019; Maslog and Lee, 2005), and (3) geopolitical actors and events which were defined inductively from the data. The coding categories are described in Table 1.

**Table 1** Framing indicators and Actor and Events Categories

	<b>Selected example of terms and phrases categorized:</b>
Conflict	<i>attack, combat, army, fighter, kill, military, missile, fleet, war, tank</i>
Human Interest	<i>boy, girl, woman, man, family, cat, dog, sex, victim, Prince Harry, Meghan Markle</i>
Morality	<i>corrupt, humanitarian, Orthodox Church, Pope Francis, Patriarch, Muslim, Islamic</i>
Responsibility	<i>accuse, alleged, accountable, provocation, responsible, denial, Kremlin comment</i>
Consequences	<i>sanction, dollar, euro, ruble, gas prices, bank, Turkish Stream, Nord Stream</i>
Crises	<i>Pandemic, Brexit, Migration, Protests, Activists, Apocalypse, Asteroid, Corruption, Climate Change, Natural Disasters, Hackers, Terrorism, Aircraft Disasters, Causes of Cancer, Conspiracy</i>
Invisible aspects of war	Peace Journalism – emotional trauma & economic loss: <i>trauma, humanitarian, white helmets</i>
Win-Win solutions	Peace Journalism – peace initiatives: <i>peace talks, accord, agreement, deal, cooperation, discussed the situation</i>
Common people as sources	Peace Journalism – statements from ordinary people: <i>man said, woman said</i>
Visible aspects of war	War Journalism – casualty counts and physical damage: <i>casualties, people died, injure, destroy</i>
Win-Loss scenarios	War Journalism – victory/defeat as end of conflict: <i>defeat, collapse, conquer, victory, win</i>
Vilification	War Journalism – demonizing language: <i>aggressor, invader, threat, attacker</i>
Victimization	War Journalism – victimizing language: <i>discriminated, defeated, unfair</i>
Emotive language	War Journalism – emotive language: <i>fear, suffer, terrible</i>

For the third research question, each of the three datasets were coded with the dominant emotion label based on Eckman's categorization (Anger, Disgust, Fear, Joy, Sadness, Surprise) using the model trained by Colnerić and Demšar (2016) on Twitter data.

### 3.2 Qualitative Approach

In the preliminary exploration of the dataset, by applying NER with Google Pinpoint, NATO was identified as the most prominent organization. For the Czech and Polish corpora, NATO was the most frequently mentioned entity ( $n_{CZ}=951$  and  $n_{PL}=1076$ ), while in the Romanian corpus, NATO was the second most named entity ( $n_{RO}=433$ , after the European Union with  $n_{EU}=1034$ ). The subcorpus for qualitative analysis was constructed by filtering the headlines containing the word "NATO" from the entire dataset (2504 headlines), out of which we selected for qualitative analysis the headlines published in 2017, 2021, and 2022 ( $n_Q=1101$ ).

The NATO subcorpus was manually coded by two coders in three steps. First, any particular definitions of the Alliance or moral evaluations pertaining to the roles, relations, and actions of NATO identified in the headlines were coded as frames by the main coder. During the second stage, the identified frames were further refined into several categories, with a discourse-mythological approach partially drawing on Moore and Gillette (1991) and Kelsey et al. (2023): NATO the Tyrant (the alliance is described as a military force, in control) and NATO the Weakling (the Alliance is described as being obsolete, weak, with negative impact); conflictual relations – building on the Cold War narrative (NATO Warrior versus Russia – where NATO is prevalent; Russia Warrior versus NATO, where Russia is prevalent); Cooperation (meetings, talks, potential joint activities between NATO and Russia); and in the context of the Russian-Ukrainian war, a frame associated with the Ukraine accession to NATO was separately coded. In the third phase, the two coders independently coded the subset of 1101 headings, and the resulting intercoder agreement score was 81.1 percent.

## 4 Findings

### 4.1 Prominent Actors and Relations

The co-occurrence analysis reveals similarities between the clustering of co-occurring actors (individual and collective). In all three cases, the two largest clusters reveal an agenda constructed around (1) US-Russia relations and conflicts with global implications in the Middle East and East Asia and (2) European issues. The clustering in the Czech and Polish corpora are more similar, but it is interesting to note that NATO clusters with the European issues nodes in the Czech case (Fig. 1) and with the US-Russia/Global Conflicts cluster in the Polish case (Fig. 2). Furthermore, in the Czech case, the Ukraine conflict clusters separately from the European issues cluster, whereas in the Polish case, it is part of it. In the Romanian case (Fig. 3), the US-Russia/Global conflict cluster is still the largest, but the European issues cluster also includes prominent occurrences of various souverainist

**Subgraph:**

01	08
02	09
03	10
04	11
05	12
06	13
07	14

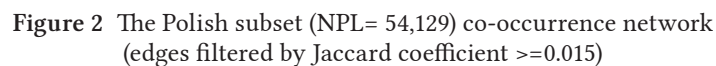
**Coefficient:**

- 0.05
- 0.10
- 0.15

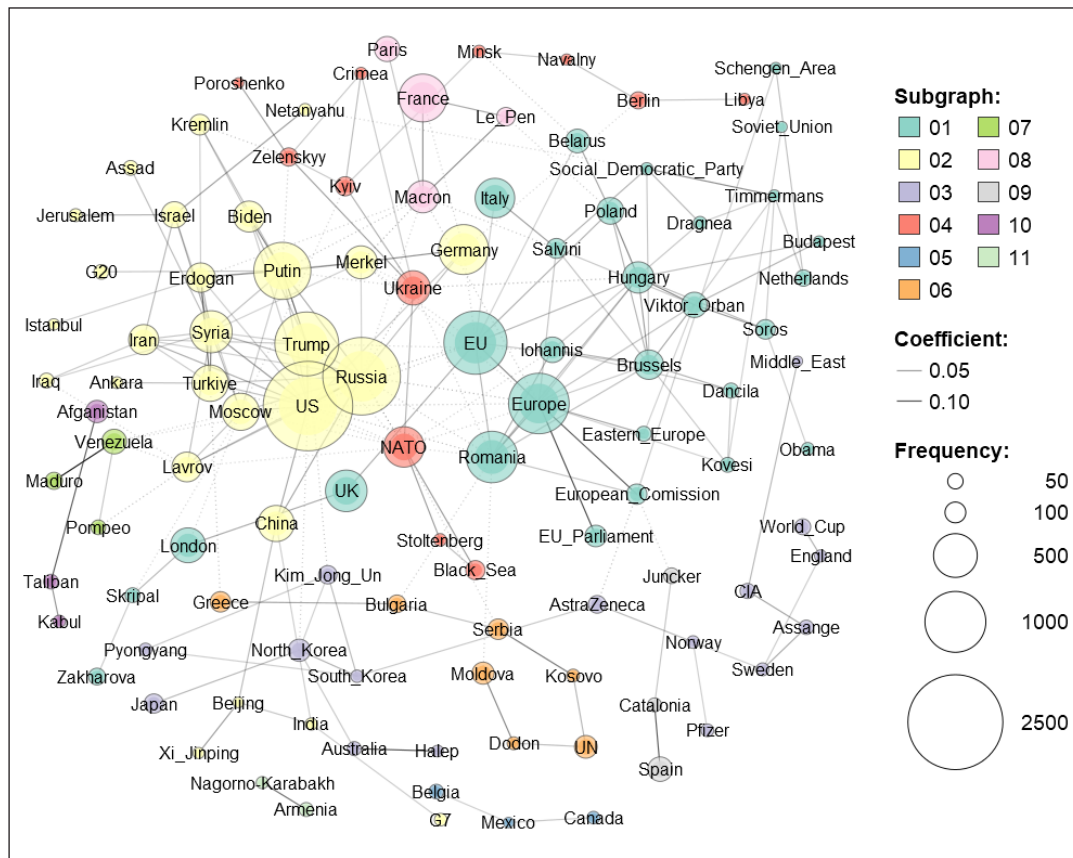
**Frequency:**

- 50
- 100
- 500
- 1000
- 2500
- 5000

INTERSECTIONS. EAST EUROPEAN JOURNAL OF SOCIETY AND POLITICS, 11(1): 63–85.



The most frequent frame is NATO the Weakling: According to this frame, NATO could have started WW3 with the USSR and is even rarely associated with Nazis (Germans during WWII); politicians from NATO member countries talk about withdrawal from NATO/or military exercises; various pundits comments about the end of NATO ('Turkey's defense minister accused NATO of failing to meet its commitments,' Sputnik CZ, 18/02/2017); The NATO-Serbia relation is presented as tense, with Serbia being against joining NATO, the enemy who bombed their country in the 1990s ('Serbs do not forgive NATO - 18 years since the attack on Belgrade,' Sputnik RO, 25/03/2017); and human interest stories about citizens of the Baltic states having negative reactions towards NATO soldiers were also associated with this frame ('Lavrov: The SS Legionnaires' March in Riga is a disgrace to Europe and NATO,' Sputnik PL, 17/03/2017).



**Figure 3** The Romanian subset (NRO= 16,136) co-occurrence network (edges filtered by Jaccard coefficient  $\geq 0.015$ )

The second and third most frequent frames represent NATO in a more active role, either as having aggressive tendencies in relation to Russia or as a ruler who seeks absolute control.

NATO Warrior vs. Russia: The headlines mention NATO forces at the Russian border or how NATO interacts with Russian aircraft or ships ('A member of the State Duma admitted the possibility of a 'credible attack' by NATO forces on Russia,' Sputnik CZ, 20/02/2017; 'NATO should fight terrorism, and not accumulate strength at the Russian border,' Sputnik PL, 18/05/2017; 'NATO again performs exercises in the Black Sea near Russian borders,' Sputnik Ro, 23/03/2021).

NATO the Tyrant frame focuses on the size of military exercises, NATO ships in the Black Sea/Odesa port, NATO acting in Afghanistan and Syria ('NATO will increase its presence on the Black Sea,' Sputnik PL, 31/01/2017; 'NATO aircraft in the sky over the EU,' Sputnik PL, 04/03/2021; 'NATO tripled the number of emergency forces,' Sputnik CZ, 17/12/2017; 'Stoltenberg: The EU cannot work without NATO,' Sputnik PL, 28/08/2017,' We will show who is the master in the house': NATO sends troops to Ukraine, Sputnik RO, 17/06/2021).



A significant number of headlines focus on Ukraine joining NATO, covering ongoing debate from both sides: NATO and Russia, and the Ukrainian army being trained and equipped by NATO and also receiving military equipment ('NATO instructors arrived at Donbas, announced in LLR,' Sputnik CZ, 15/09/2017). Ukraine in NATO is negatively framed as the headlines emphasize that joining NATO will not solve the problem in Donbas ('Ukraine in NATO? This is the worst decision,' Sputnik PL, 10/07/2017; 'German expert: Why Ukraine and Georgia will not be able to join NATO,' Sputnik RO, 24/03/2021).

A frame less frequently found attributes a more active role to Russia with respect to NATO, mainly through the use of key quotes from belligerent Russian sources. Russia Warrior vs. NATO relies on Putin's, Zakharchenko's, Shoigu's, and Lavrov's statements about NATO and official statements attributed to The Kremlin, Russian Federation, and Ministry of Defense ('Putin: Russia will not be 'passively looking' for NATO expansion,' Sputnik PL, 1/06/2017; 'Diplomat: Russian Navy is a real threat to NATO,' Sputnik CZ, 13/02/2021; 'Russian hunting plane intercepts a new NATO plane close to Russia's borders,' Sputnik RO, 26/05/2021).

The least prominent frame is that of Cooperation NATO – Russia. This focuses on dialogue, councils, meetings, or statements about Russia joining NATO ('Stoltenberg: NATO needs a dialogue with Russia, not a new Cold War,' Sputnik PL, 25/05/2017; 'Russia and NATO at the shared table,' Sputnik PL, 13/07/2017; 'The Russian and NATO army participate in common exercises for the first time in 10 years,' Sputnik RO, 11/02/2021, 'The Russian-NATO Council started in Brussels,' Sputnik CZ, 12/01/2021).

Table 7 Distribution of coded frames by edition

FRAME	CZECH	POLISH	ROMANIAN	Total
NATO Weakling	118	122	72	312
NATO Warrior vs Russia	108	67	26	201
NATO Tyrant	67	91	35	193
Ukraine in NATO	73	87	15	175
Russia Warrior vs NATO	55	53	28	136
Cooperation NATO – Russia	34	31	19	84

As Table 7 and Fig. 6 show, although NATO the Weakling frame is dominant in all editions and both timeframes analyzed, the two frames that attribute an aggressive or controlling role are also significantly present in all editions. The decrease between 2017 and 2021 is due to a gradual decline in the number of articles published by Sputnik in the international sections to almost a third. Of the two frames that attribute an active role to NATO, in the Czech edition, NATO Warrior vs Russia is more frequently employed, whereas in the Polish edition, NATO the Tyrant is more frequently employed. In the Romanian edition, NATO the Tyrant emerges as the more dominant of the two frames in 2021.



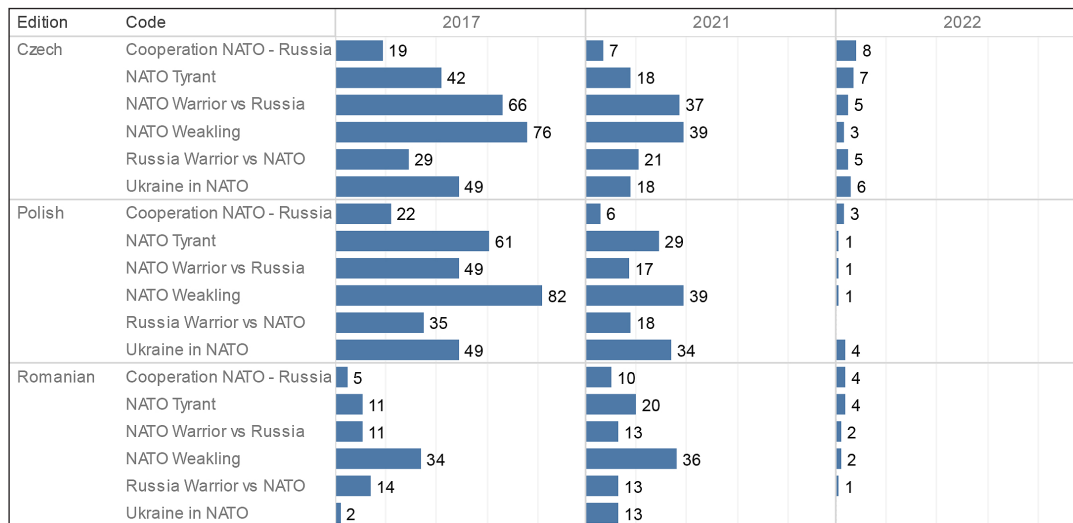


Figure 6 Overview of manually coded framings of headlines mentioning NATO by year

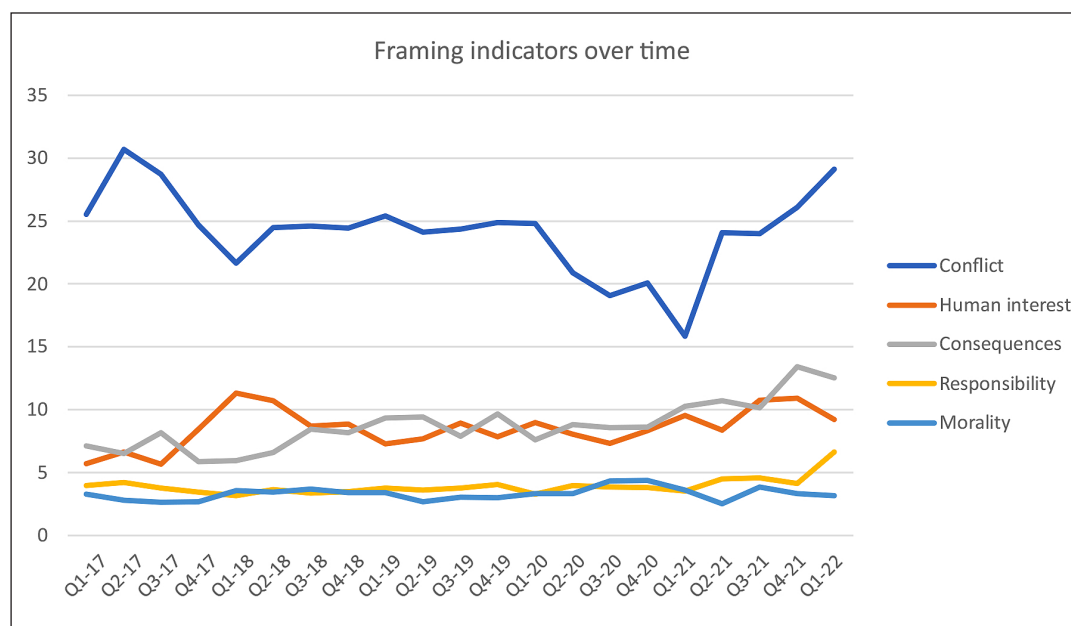
#### 4.3 Framing and Peace-War Journalism Coverage Indicators

The coding of generic frame indicators reveals an overwhelming focus on conflict, with almost a quarter of total headlines featuring indicators of the frame. Human interest is the second most prominent, but it is worth noting that the Czech edition features significantly more headlines coded as human interest (news about the British Royal Family – particularly focusing on Meghan Markle and Prince Harry, Game of Thrones TV show news, and clickbait articles focused on sex, which do not appear in the other two editions). The third most prominent group of framing indicators are (economic) consequences, where references to Nord Stream 2 and other gas pipeline projects were coded besides references to costs, millions or billions of euros or dollars. The Czech edition seems to be the only one that slightly differs in terms of the distribution of framing indicators, with a greater focus on human interest and slightly more occurrences of indicators of attribution of responsibility.

Table 2 Overview of generic framing indicators coded on English translated full dataset (N=118,198)

Framing indicators	No. Cases	% Total cases	% Polish cases	% Czech cases	% Romanian cases
Conflict	28740	24.54%	23.32%	27.15%	20.97%
Human interest	9748	8.32%	6.93%	10.14%	7.66%
Consequences	9627	8.22%	7.44%	8.96%	8.67%
Morality	3995	3.41%	3.13%	3.78%	3.24%
Responsibility	4400	3.76%	3.28%	4.36%	3.57%

Over time, the overall number of articles published in the international news section steadily declined throughout the timeframe, from 30,315 in 2017 to 15,133 in 2021. The January 2022 data only amounts to 694 headlines. However, as Fig. 4 shows, according to the proportion of cases with respect to total headlines published, there was a decrease in conflict frame indicators during the COVID-19 pandemic, followed by a steep increase in 2021. The (economic) consequences and attribution of responsibility framing indicators also seem to register an increase in 2021, especially in the first month of 2022.



**Figure 4** Framing indicators (as percentage of total headlines) change over time (N=118,198)

The most prominent indicators for War Journalism / Peace Journalism coverage use the elites as sources, with frequent reporting on statements made by Putin and Trump, as well as various ministries and Russian public figures (such as Maria Zakharova). Mentions of win-win finalities (peace talks, accords, discussions) and win-loss scenarios (defeat, victory) also appear relatively frequently. It is interesting to note that win-win peace journalism coverage clusters with the economic consequences and attribution of responsibility frame indicators (Fig. 5), while the win-loss war journalism coverage clusters with the morality frame indicators, and both also cluster with indicators of emotive language. The conflict frame indicators cluster with indicators of war journalism coverage, such as the vilification of the enemy and the prominent use of elite sources. The cluster analysis of codes in Figure 5 suggests several discursive strategies, listed here in order of their prominence: (1) the vilification-oriented use of statements and key quotes in conflict-focused news

coverage of US-Russia and global conflict; (2) crisis- (Western Europe) and dissent- (Eastern Europe) oriented news with respect to Europe and the EU; (3) the negotiation-oriented use of economic consequences and attribution of responsibility frames in the context of peace journalism coverage mentioning win-win solutions; (4) moral-positioning-oriented emotive coverage of win-loss scenarios.

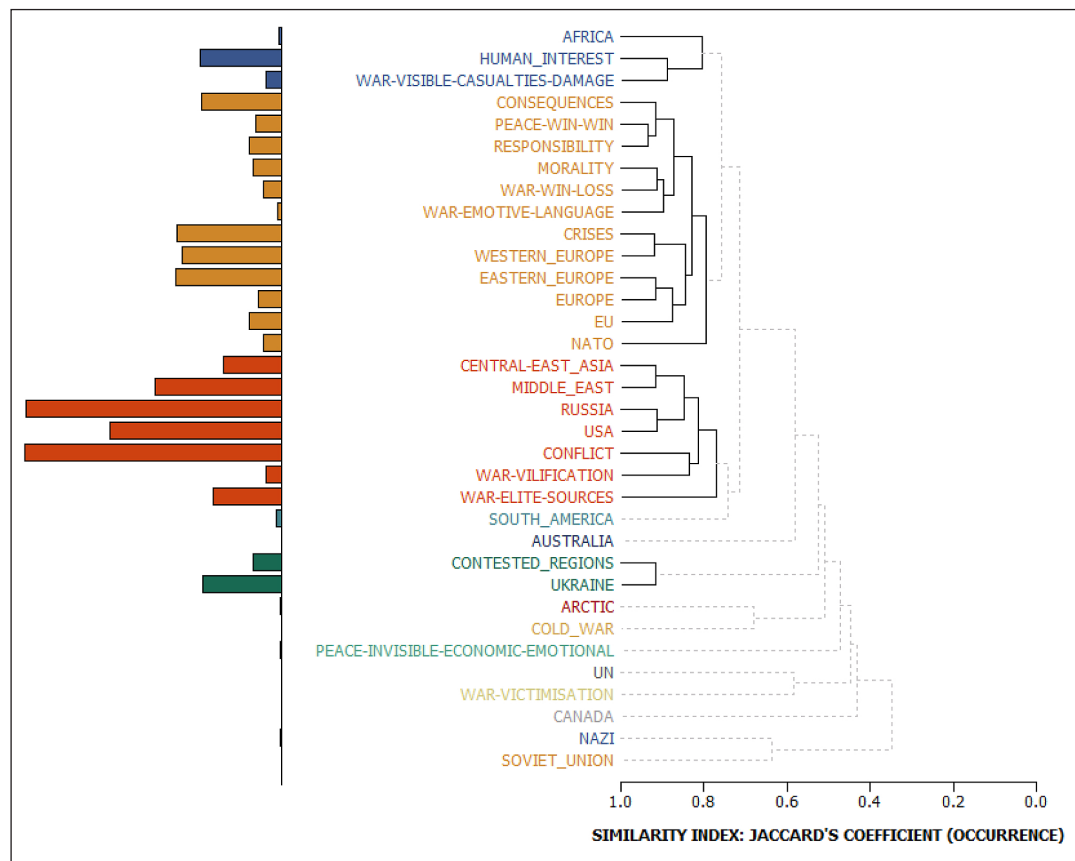


Figure 5 Clustering of framing indicators, war-peace coverage orientation indicators and prominent issues and actors by co-occurrence

#### 4.4 Affective framing and prominent individual actors

As affective framing seems prominently employed, especially in the context of elite actors/sources, the categorization of emotion provides insight into the most frequent dominant emotion. Table 3 reveals that the most likely dominant emotion in the Czech and Polish cases is Fear, followed by Surprise and Joy. The same three emotions were categorized as the most prominent in the Romanian case as well, but Surprise is more prominent than

Fear. In the Czech case, Fear and Surprise are present in almost the same overall proportions. The emphasis on news values such as negativity and novelty probably plays a role in the two most frequently categorized emotions. Anger and Disgust are the least frequent categories identified.

In the analysis of the mentions of the most prominent individual actors with respect to emotional categorization we found that, in the three editions, Trump is the individual most associated with detected negative emotions such as with Anger and Disgust out of the small number of such cases. Putin and Trump are the two figures who are most frequently mentioned in all three editions. In all three editions, the two appear in a significant number of cases where Fear was detected as the dominant emotion. Both in terms of the most prominent individuals and in terms of emotional patterns, the Czech and Polish editions are similar, while the Romanian edition also includes Western European actors such as Merkel and Macron, but also souverainists such as Le Pen, Salvini and Orbán, and conspiracy theory scapegoats such as Soros.

**Table 3** Categorization by dominant emotion of headlines  
(machine translated into English)

Emotion Categorized	CZECH Headlines (%)	POLISH Headlines (%)	ROMANIAN Headlines (%)
Anger	746 (1.59%)	715 (1.32%)	199 (1.23%)
Disgust	314 (0.67%)	192 (0.35%)	68 (0.42%)
Fear	17251 (36.80%)	18553 (34.28%)	5136 (31.83%)
Joy	9652 (20.59%)	15010 (27.73%)	4295 (26.62%)
Sadness	1897 (4.05%)	2931 (5.41%)	834 (5.17%)
Surprise	17013 (36.30%)	16728 (30.90%)	5604 (34.73%)
Total	46873	54129	16136

Table 4 Emotional categorization (automated) with respect to mentions of most prominent individuals<sup>2</sup>

CZECH (N <sub>CZ</sub> = 46873) Actor/Emotion	Anger	Disgust	Fear	Joy	Sadness	Surprise	Total	chi-square
Trump	71 (9.52%)	15 (4.78%)	561 (3.25%)	333 (3.45%)	55 (2.90%)	541 (3.18%)	1576 (3.36%)	92.785**
Putin	15 (2.01%)	5 (1.59%)	501 (2.90%)	377 (3.91%)	29 (1.53%)	617 (3.63%)	1544 (3.29%)	50.753**
Lavrov	8 (1.07%)	0 (0.00%)	182 (1.06%)	170 (1.76%)	14 (0.74%)	102 (0.60%)	476 (1.02%)	87.658**
Poroshenko	3 (0.40%)	1 (0.32%)	117 (0.68%)	72 (0.75%)	14 (0.74%)	118 (0.69%)	325 (0.69%)	2.06
Erdogan	3 (0.40%)	5 (1.59%)	90 (0.52%)	24 (0.25%)	11 (0.58%)	104 (0.61%)	237 (0.51%)	24.273**
Merkel	3 (0.40%)	0 (0.00%)	63 (0.37%)	60 (0.62%)	8 (0.42%)	83 (0.49%)	217 (0.46%)	10.671
Zakharova	1 (0.13%)	2 (0.64%)	82 (0.48%)	21 (0.22%)	2 (0.11%)	100 (0.59%)	208 (0.44%)	26.355**
Biden	3 (0.40%)	2 (0.64%)	48 (0.28%)	28 (0.29%)	4 (0.21%)	109 (0.64%)	194 (0.41%)	34.800**
M. Markle	9 (1.21%)	4 (1.27%)	31 (0.18%)	29 (0.30%)	15 (0.79%)	98 (0.58%)	186 (0.40%)	62.598**
Macron	3 (0.40%)	1 (0.32%)	48 (0.28%)	38 (0.39%)	7 (0.37%)	87 (0.51%)	184 (0.39%)	11.981*
Peskov	1 (0.13%)	0 (0.00%)	54 (0.31%)	39 (0.40%)	6 (0.32%)	65 (0.38%)	165 (0.35%)	4.119
Navalny	9 (1.21%)	1 (0.32%)	55 (0.32%)	31 (0.32%)	4 (0.21%)	64 (0.38%)	164 (0.35%)	17.802**
Lukashenko	3 (0.40%)	1 (0.32%)	56 (0.32%)	18 (0.19%)	4 (0.21%)	69 (0.41%)	151 (0.32%)	10.103
Assad	5 (0.67%)	1 (0.32%)	75 (0.43%)	13 (0.13%)	34 (1.79%)	18 (0.11%)	146 (0.31%)	178.397**
Kim Jong Un	1 (0.13%)	0 (0.00%)	17 (0.10%)	27 (0.28%)	6 (0.32%)	60 (0.35%)	111 (0.24%)	25.965**

<sup>2</sup> In Tables 4, 5, and 6, percentages in parentheses indicate the proportion of headlines categorized with each emotion which contain references to the respective actor. Totals for each categorized emotion by subset are shown in Table 3. Two asterisks indicate an alpha level of 0.01 (i.e.,  $p < 0.01$ ), while one asterisk indicates an alpha level of 0.05. The chi-squared test is used to assess the association between the mentioned actor and categorized emotion. For each actor, values exceeding 1 percent (of headlines categorized with the respective emotion) are highlighted in the tables.

Table 5 Emotional categorization (automated) with respect to mentions of most prominent individuals<sup>3</sup>

POLISH (N <sub>PL</sub> = 54129) Actor/Emotion	Anger	Disgust	Fear	Joy	Sadness	Surprise	Total	chi-square
Trump	45 (6.29%)	12 (6.25%)	600 (3.23%)	561 (3.74%)	127 (4.33%)	504 (3.01%)	1849 (3.42%)	44.895**
Putin	11 (1.54%)	1 (0.52%)	457 (2.46%)	553 (3.68%)	56 (1.91%)	506 (3.02%)	1584 (2.93%)	64.333**
Poroshenko	6 (0.84%)	1 (0.52%)	93 (0.50%)	169 (1.13%)	16 (0.55%)	195 (1.17%)	480 (0.89%)	60.133**
Lukashenko	3 (0.42%)	2 (1.04%)	141 (0.76%)	65 (0.43%)	18 (0.61%)	164 (0.98%)	393 (0.73%)	34.895**
Lavrov	11 (1.54%)	3 (1.56%)	146 (0.79%)	91 (0.61%)	26 (0.89%)	99 (0.59%)	376 (0.69%)	17.604**
Navalny	3 (0.42%)	2 (1.04%)	124 (0.67%)	78 (0.52%)	10 (0.34%)	83 (0.50%)	300 (0.55%)	9.21
Zelenskyy	3 (0.42%)	1 (0.52%)	94 (0.51%)	80 (0.53%)	14 (0.48%)	102 (0.61%)	294 (0.54%)	2.297
Merkel	6 (0.84%)	0 (0.00%)	60 (0.32%)	101 (0.67%)	16 (0.55%)	78 (0.47%)	261 (0.48%)	24.289**
Biden	4 (0.56%)	2 (1.04%)	74 (0.40%)	32 (0.21%)	8 (0.27%)	134 (0.80%)	254 (0.47%)	66.361**
Erdogan	3 (0.42%)	1 (0.52%)	79 (0.43%)	41 (0.27%)	28 (0.96%)	89 (0.53%)	241 (0.45%)	30.269**
Macron	5 (0.70%)	1 (0.52%)	63 (0.34%)	63 (0.42%)	11 (0.38%)	82 (0.49%)	225 (0.42%)	6.402
Kim Jong Un	1 (0.14%)	0 (0.00%)	40 (0.22%)	45 (0.30%)	7 (0.24%)	105 (0.63%)	198 (0.37%)	47.762**
Duda	1 (0.14%)	0 (0.00%)	49 (0.26%)	37 (0.25%)	13 (0.44%)	60 (0.36%)	160 (0.30%)	7.445
Saakashvili	0 (0.00%)	0 (0.00%)	29 (0.16%)	67 (0.45%)	8 (0.27%)	52 (0.31%)	156 (0.29%)	27.242**
Zakharova	2 (0.28%)	4 (2.08%)	44 (0.24%)	14 (0.09%)	8 (0.27%)	76 (0.45%)	148 (0.27%)	61.905**

<sup>3</sup> In Tables 4, 5, and 6, percentages in parentheses indicate the proportion of headlines categorized with each emotion which contain references to the respective actor. Totals for each categorized emotion by subset are shown in Table 3. Two asterisks indicate an alpha level of 0.01 (i.e.,  $p < 0.01$ ), while one asterisk indicates an alpha level of 0.05. The chi-squared test is used to assess the association between the mentioned actor and categorized emotion. For each actor, values exceeding 1 percent (of headlines categorized with the respective emotion) are highlighted in the tables.

Table 6 Emotional categorization (automated) with respect to mentions of most prominent individuals<sup>4</sup>

ROMANIAN (N <sub>RO</sub> = 16136) Actor/Emotion	Anger	Disgust	Fear	Joy	Sadness	Surprise	Total	chi-square
Trump	30 (15.08%)	5 (7.35%)	352 (6.85%)	314 (7.31%)	58 (6.95%)	348 (6.21%)	1107 (6.86%)	26.131**
Putin	12 (6.03%)	4 (5.88%)	231 (4.50%)	298 (6.94%)	30 (3.60%)	302 (5.39%)	877 (5.44%)	33.333**
Merkel	4 (2.01%)	0 (0.00%)	51 (0.99%)	94 (2.19%)	11 (1.32%)	113 (2.02%)	273 (1.69%)	26.992**
Macron	3 (1.51%)	0 (0.00%)	63 (1.23%)	61 (1.42%)	21 (2.52%)	111 (1.98%)	259 (1.61%)	16.116**
Biden	3 (1.51%)	1 (1.47%)	61 (1.19%)	62 (1.44%)	3 (0.36%)	111 (1.98%)	241 (1.49%)	19.667**
Lavrov	2 (1.01%)	1 (1.47%)	84 (1.64%)	78 (1.82%)	8 (0.96%)	48 (0.86%)	221 (1.37%)	21.187**
Erdogan	8 (4.02%)	3 (4.41%)	92 (1.79%)	29 (0.68%)	9 (1.08%)	73 (1.30%)	214 (1.33%)	38.794**
Iohannis	2 (1.01%)	1 (1.47%)	44 (0.86%)	54 (1.26%)	5 (0.60%)	52 (0.93%)	158 (0.98%)	5.783
Viktor Orbán	4 (2.01%)	0 (0.00%)	48 (0.93%)	40 (0.93%)	4 (0.48%)	59 (1.05%)	155 (0.96%)	5.568
Salvini	0 (0.00%)	0 (0.00%)	30 (0.58%)	32 (0.75%)	4 (0.48%)	25 (0.45%)	91 (0.56%)	5.557
Soros	1 (0.50%)	0 (0.00%)	35 (0.68%)	21 (0.49%)	3 (0.36%)	27 (0.48%)	87 (0.54%)	3.36
Le Pen	0 (0.00%)	0 (0.00%)	19 (0.37%)	17 (0.40%)	4 (0.48%)	36 (0.64%)	76 (0.47%)	6.414
Kim Jong Un	1 (0.50%)	0 (0.00%)	10 (0.19%)	23 (0.54%)	0 (0.00%)	41 (0.73%)	75 (0.46%)	21.405**
Zelenskyy	2 (1.01%)	0 (0.00%)	19 (0.37%)	30 (0.70%)	2 (0.24%)	21 (0.37%)	74 (0.46%)	9.652
Maduro	2 (1.01%)	0 (0.00%)	14 (0.27%)	20 (0.47%)	1 (0.12%)	27 (0.48%)	64 (0.40%)	7.299

<sup>4</sup> In Tables 4, 5, and 6, percentages in parentheses indicate the proportion of headlines categorized with each emotion which contain references to the respective actor. Totals for each categorized emotion by subset are shown in Table 3. Two asterisks indicate an alpha level of 0.01 (i.e.,  $p < 0.01$ ), while one asterisk indicates an alpha level of 0.05. The chi-squared test is used to assess the association between the mentioned actor and categorized emotion. For each actor, values exceeding 1 percent (of headlines categorized with the respective emotion) are highlighted in the tables.



## 5 Discussion and Conclusion

This mixed methods analysis of a large dataset of Sputnik News headlines from three different language editions reveals both overarching discursive patterns and edition-level particularities. Sputnik's coverage of international news is dominated by indicators of conflict framing. Reflecting the role of Sputnik as an instrument of Russian foreign policy, the media propose that the United States and Russia are the main actors in the conflict, a discursive pattern that builds on and is consistent with the Cold War narratives that have defined the main poles of global power for decades. However, other prominent actors emerge in the Kremlin's more recent strategic communication, NATO and the EU, and thus the antagonism between good and evil is developed within a multidimensional ideological construction. NATO is the most frequently mentioned organization in the headlines. Considering that the Sputnik News network was created in the aftermath of the annexation of Crimea and has been constantly preoccupied with the situation of Ukraine and the Western orientation, emphasized by the country's intention to join NATO, it is not surprising that the representation of Ukraine in the sampled headlines of the Sputnik versions from the three EU and NATO member countries (Poland, Czech Republic, and Romania) is so prominent. The Czech and Polish emerging international news coverage patterns are relatively similar, with the Czech Sputnik News including a significant human-interest component not found in the other editions. The Romanian edition differs as it is more focused on the EU than on international politics and includes references to European politicians that are associated with souveranist discourses (Orbán, Le Pen, Salvini). Furthermore, international news includes more references to national politics than in the other two cases.

By far the most frequently mentioned individual actors are (former) POTUS Donald Trump and Russian Federation President Vladimir Putin. Different patterns are found with the most frequently mentioned individuals in Czech and Polish coverage in relation to the Romanian coverage. The former two include more occurrences of regional, Eastern-European actors, while the latter includes more Western European actors.

The patterns of coverage correspond to the findings of previous literature on the topic. The niche dictionary-based analysis of framing and war/peace journalism coverage indicators reveals four overarching discourses: (1) vilification associated with global conflict (US and NATO); (2) crisis and dissent (in Europe and the EU); (3) the possibility of agreement; and (4) occupying the moral high-ground. Although these are not far removed from the dismiss-distort-distract-dismay Soviet propaganda model (Snegovaya, 2015), the focus on Trump, and also (in the Romanian case) Orbán, Le Pen, Salvini, on the one hand, and George Soros, on the other, seems to favor Watanabe's (2018) conclusion that Sputnik News might play a role in perpetuating and enforcing conspiracy theories and the souveranist ideology in Europe. From this reading, even the Czech focus on the Meghan Markle case can be seen as adding fuel to the fire of moral panic about the disintegration of national European cultural reference points (such as the British Royal Family). Within the 2020–2021 timeframe, the indicators of the conflict frame decrease in frequency, but the COVID-19 pandemic and subsequent vaccination initiatives are covered and include significant conspiracy theory elements (Mogoş et al., 2022).

Furthermore, formulas that leverage human interest (death, sex, and popular TV shows or sports competitions) seem geared toward attracting readership from social media contexts, relying on 'spreadability.' The emotionalization of international coverage relies

on the voices of individual actors, most prominently Donald Trump and Vladimir Putin, with key quotes included in the headlines, allowing for more colorful, emotive language and the vilification or disparaging of other actors – the USA, NATO, the EU or European leaders. Prominent Russian voices such as Lavrov, Zakharova, and Shoigu supplement this discursive approach. Mentions of Assad and Syria, Maduro and the Venezuelan situation, or Kim Jong Un and North Korea may serve as exemplars of the possible negative outcomes of international politics. The frame analysis of the coverage of NATO reveals a tendency to depict NATO as the more active (in Warrior or Tyrant roles) counterpart of Russia, which mostly reacts to shows of force or attempts to expand or consolidate control. However, NATO is typically framed as a failing alliance with dissenting members or as lacking the coordination necessary to take decisive, successful action.

The findings of the quantitative and qualitative analysis of Sputnik News headlines fit the characteristics of the emergent ‘propaganda on demand’ strategy model put forward by Litvinenko (2022), suggesting that, similarly to the internal Russian media landscape (Litvinenko & Toepfl, 2019), the Sputnik News network employs a content strategy that is not reliant on a coherent narrative, but rather on large volumes of emotionalized eclectic messages. This amounts to contradictory narratives that are difficult to counter and that can be replicated by other news sources or shared on digital platforms where audience members may select convenient messages to support ‘post-truth’ stances.

This research contributes to the scientific literature on the topic of instrumental international news coverage by government-funded agencies through a comparative analysis of three language editions. The methodological approach combines quantitative methodologies using a large multilingual dataset and Named Entity Recognition, niche dictionaries, the inductive and deductive coding of frame indicators, content analysis, and qualitative analysis on a filtered subset. A limitation of the research is the use of automated translations for some parts of the analysis. With the development of more natural language processing tools for languages such as Czech, Polish, and Romanian, this can be overcome in further analyses. Further research is also needed to explore the representation of other prominent actors like the EU, as well as the domestication strategies employed in each of the three editions.

## References

- Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analyzing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, 25(2), 135–158.
- Caple, H., & Bednarek, M. (2016). Rethinking news values: What a discursive approach can tell us about the construction of news discourse and news photography. *Journalism*, 17(4), 435–455.
- Colnerič, N., & Demšar, J. (2018). Emotion recognition on Twitter: Comparative study and training a unison model. *IEEE transactions on affective computing*, 11(3), 433–446.
- Demšar, J., Curk, T., Erjavec, A., Gorup, Č., Hočevár, T., Milutinovič, M., ... & Zupan, B. (2013). Orange: data mining toolbox in Python. *the Journal of machine Learning research*, 14(1), 2349–2353.

- Demjanski, A. (2020). *Framing the news: Deutsche Welle, Sputnik News, and Macedonia's Colorful Revolution*. [Unpublished Master dissertation]. University of Texas at Austin. <https://bit.ly/3q5BPmR>
- Deverell, E., Wagnsson, C., & Olsson, E.K. (2021). Destruct, direct and suppress: Sputnik narratives on the Nordic countries. *Journal of International Communication*, 27(1), 15–37. <https://doi.org/10.1080/13216597.2020.1817122>
- Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion*, 6(3-4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), 124. <https://doi.org/10.1037/h0030377>
- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51–58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- EUvsDisinfo. (2019). *Figure of the Week: 1.3 Billion – EUvsDisinfo*. <https://euvsdisinfo.eu/figure-of-the-week-1-3-billion> Accessed: 01.08.2023.
- Furman, I. O., Gürel, K. B., & Sivaslıoğlu, F. B. (2023). 'As Reliable as a Kalashnikov Rifle': How Sputnik News Promotes Russian Vaccine Technologies in the Turkish Twittersphere. *Social Media+ Society*, 9(1), 20563051221150418. <https://doi.org/10.1177/20563051221150418>
- Guo, L., Su, C., Paik, S., Bhatia, V., Akavoor, V. P., Gao, G., ... & Wijaya, D. (2023). Proposing an open-sourced tool for computational framing analysis of multilingual data. *Digital journalism*, 11(2), 276–297. <https://doi.org/10.1080/21670811.2022.2031241>
- Herman, L., & Vervaeck, B. (2019). *Handbook of narrative analysis*. University of Nebraska Press.
- Higuchi, K. (2016). KH Coder (Version 3). Ritsumeikan University. URL: <https://kxhcodr.net/en/>
- Jaitner, M., & Mattsson, P. A. (2015, May). Russian information warfare of 2014. In *2015 7th International Conference on Cyber Conflict: Architectures in Cyberspace* (pp. 39–52). IEEE.
- Kelsey, D., Yannopoulou, N., Whittle, A., Heath, T., Golossenko, A., & Soares, A. M. (2023). The (army) hero with a thousand faces: A discourse-mythological approach to theorising archetypal blending in contemporary advertising. *Marketing Theory*, 23(1), 141–162. <https://doi.org/10.1177/14705931221141732>
- Kragh, M., & Åsberg, S. (2017). Russia's strategy for influence through public diplomacy and active measures: the Swedish case. *Journal of Strategic Studies*, 40(6), 773–816. <https://doi.org/10.1080/01402390.2016.1273830>
- Kuusisto, R. (2018). Comparing IR plots: dismal tragedies, exuberant romances, hopeful comedies and cynical satires. *International Politics*, 55, 160–176. <https://doi.org/10.1057/s41311-017-0076-2>
- Lams, L., De Smaele, H., De Coninck, F., Lippens, C., & Smeyers, L. (2023). Strategic Comrades? Russian and Chinese Media Representations of NATO. *Europe-Asia Studies*, 75(5), 842–864. <https://doi.org/10.1080/09668136.2022.2152556>
- Litvinenko, A. (2022). Propaganda on demand: Russia's media environment during the war in Ukraine. *Global Media Journal – German Edition*, 12(2), 1–14. <https://doi.org/10.22032/dbt.55518>
- Litvinenko, A., & Toepfl, F. (2019). The 'Gardening' of an authoritarian public at large: How Russia's ruling elites transformed the country's media landscape after the 2011/12 protests 'For Fair Elections'. *Publizistik*, 64, 225–240. <https://doi.org/10.1007/s11616-019-00486-2>
- Maslog, C., & Lee, S. (2005). War or peace journalism? Asian newspaper coverage of conflicts. *Journal of Communication*, 55(2), 311–329. <https://doi.org/10.1111/j.1460-2466.2005.tb02674.x>

- Mogoș, A. A., Grapă, T. E., & Șandru, T. F. (2022). Russian disinformation in Eastern Europe. Vaccination media frames in ro.sputnik.md. *Comunicar*, 30(72), 33–46. <https://doi.org/10.3916/C72-2022-03>
- Moore, R. & Gillette, D. (1991). *King, Warrior, Magician, Lover: Rediscovering the Archetypes of the Mature Masculine*, HarperOne.
- Müller, P., & Schulz, A. (2021). Alternative media for a populist audience? Exploring political and media use predictors of exposure to Breitbart, Sputnik, and Co. *Information, Communication & Society*, 24(2), 277–293. <https://doi.org/10.1080/1369118X.2019.1646778>
- Nelson, T. S. (2019). How RT frames conflict: A comparative analysis. *Russian Journal of Communication*, 11(2), 126–140. <https://doi.org/10.1080/19409419.2019.1608846>
- Robertson, A. (2017). Narrative analysis. In K. Boréus. & G. Bergström (Eds.), *Analyzing text and discourse: Eight approaches for the social sciences*, (pp. 122–145). Sage Publications.
- Semetko, H.A., & Valkenburg, P.M. (2000). Framing European politics: A content analysis of press and television news. *Journal of Communication*, 50(2), 93–109. <https://doi.org/10.1111/j.1460-2466.2000.tb02843.x>
- Shahzad, F., Qazi, T. A., & Shehzad, R. (2023). Framing of Israel and Palestine Conflict in RT news, Al-Jazeera, CNN & BBC News. *Global Digital & Print Media Review*, 6(2), 1–14. [https://doi.org/10.31703/gdpmr.2023\(VI-II\).01](https://doi.org/10.31703/gdpmr.2023(VI-II).01)
- Snegovaya, M. (2015). Putin's information warfare in Ukraine. *Soviet Origins of Russia's Hybrid Warfare*, *Russia Report*, 1, 133–135.
- Ștefan, B. (2020). Understanding Sputnik news agency internet traffic analysis. *Bulletin of the Transilvania University of Brasov*, 13(1), 113–124.
- US Department of State Global Engagement Center. (2022). *Kremlin-Funded Media: RT and Sputnik's Role in Russia's Disinformation and Propaganda Ecosystem*. [https://www.state.gov/wp-content/uploads/2022/01/Kremlin-Funded-Media\\_January\\_update-19.pdf](https://www.state.gov/wp-content/uploads/2022/01/Kremlin-Funded-Media_January_update-19.pdf) Accessed: 01.08.2023.
- Toepfl, F. (2020). Comparing authoritarian publics: The benefits and risks of three types of publics for autocrats. *Communication Theory*, 30(2), 105–125. <https://doi.org/10.1093/ct/qtz015>
- Van Dijk, T. A. (1993). Van Dijk, T. A. (1993). Stories and Racism. In D. K. Mumby (Ed.) *Narrative and Social Control: Critical Perspectives*, (pp. 121–142). Sage Publications.
- Watanabe, K. (2018, August). *Conspiracist propaganda: How Russia promotes anti-establishment sentiment online*. In ECPR General Conference, Hamburg, Germany.
- Wahl-Jorgensen, K. (2020). An emotional turn in journalism studies? *Digital Journalism*, 8(2), 175–194.
- Wahl-Jorgensen, K. (2019). *Emotions, media and politics*. John Wiley & Sons. <https://doi.org/10.1080/21670811.2019.1697626>
- White, H. (1973). *Metahistory: The historical imagination in nineteenth-century Europe*. John Hopkins University Press.
- Wilbur, D. S. (2021). Propaganda or Not. *Journal of Information Warfare*, 20(3), 146–156. <https://www.jstor.org/stable/27125004>
- Wilson, J. L. (2015). Russia and China respond to soft power: Interpretation and readaptation of a Western construct. *Politics*, 35(3-4), pp. 287–300. <https://doi.org/10.1111/1467-9256.12095>
- Yablokov, I. (2015). Conspiracy theories as a Russian public diplomacy tool: The case of Russia Today (RT). *Politics*, 35(3-4), 301–315. <https://doi.org/10.1111/1467-9256.12097>

---

ILYA SULZHYTSKI,\* & VARVARA KULHAYEVA\*\*

## Ambassadors of War: Social and Semantic Networks of Belarusian Pro-Government Telegram Channels during the 2022 Russian Invasion of Ukraine

Intersections. EEJSP

11(1): 86–110.

<https://doi.org/10.17356/ieejsp.v11i1.1224>

<https://intersections.tk.hu>

---

\* [ilya.sulzhickiy@gmail.com](mailto:ilya.sulzhickiy@gmail.com) (Institute of Slavic Studies, University of Innsbruck, Innsbruck, Austria)

\*\* [v.a.kulgaeva@gmail.com](mailto:v.a.kulgaeva@gmail.com) (independent researcher)

### Abstract

This study examined the response of Belarusian pro-government Telegram channels to the 2022 Russian invasion of Ukraine in its initial phase (22 February to 24 March 2022). A socio-semantic network approach was used to analyse the relationships between 21 influential pro-government channels and the concepts they disseminate. We applied a four-stage framework based on mutual citation analysis, extended hyperlink analysis, bipartite channel-concept network analysis, and inter-conceptual relations analysis.

The findings indicate that Belarusian pro-government channels function as critical intermediaries between Russian media and local Lukashenka supporters. The network structure reveals two principal centres of media activity: one focusing on disseminating pro-Russian military content, and the other propagating messages supporting the Belarusian government. This distinction is also evident in the main topics discussed, which can be categorised as pro-Russian warfare, pro-Russian foreign policy, Belarusian domestic policy, and pro-Lukashenka content. Notably, the high proportion of war-related concepts within this semantic structure suggests that Russian political language has significantly influenced the network of pro-government channels, challenging the maintenance of domestic Belarusian narratives. Furthermore, anonymous channels on the network's periphery disseminate more explicitly aggressive war-related concepts, potentially operating with greater autonomy from direct government control.

This study contributes to understanding information warfare dynamics in Eastern Europe during the 2022 Russian invasion of Ukraine. It provides insights into the complexities of information networks in contested political spaces by revealing how both social and semantic relations between Telegram channels shape media agendas.

**Keywords:** Belarus, Telegram, Russia, Ukraine, war, media, network analysis, 2022

## 1 Introduction

The rule of authoritarian and hybrid regimes is based on more than 'hard' power: it is impossible to rule society through repressive means and physical coercion alone. In Russia, for example, 'soft' power is actively used to control public opinion and the media agenda

---



(Mejias & Vokuev, 2017; Karpchuk & Yuskiv, 2021; Kizlova & Norris, 2022). In Belarus, the brutal repression of peaceful protests was accompanied by active media support from pro-state official and grassroots Telegram channels, YouTube bloggers, and public figures, which contributed to the escalation of conflict and polarisation of society (Navumau, 2020; Polovyi, 2021; Greene, 2022; Sulzhytski, 2022; Kuznetsova, 2023; Deikalo, 2023).

After the Russian invasion of Ukraine in 2022, the most innovative OSINT (Open-Source Intelligence) and research projects devoted all their efforts to identifying and countering the influence of pro-Russian propaganda, especially in EU countries and the US (Oates et al., 2022; Geissler et al., 2023; Jarynowski, 2022; Lawriwsky, 2023; Greene, 2023). However, these projects focus primarily on traditional sources of pro-Russian influence: Twitter (Chen & Ferrara, 2023), official media such as Russia Today (Metzger, 2023), and well-known Telegram channels in Russia and Ukraine (Nazaruk, 2022; Ptaszek et al., 2023). By contrast, a significantly smaller number of studies and projects on anti-Ukrainian propaganda highlight that supporters of the Lukashenka regime actively used Telegram to spread anti-Ukrainian propaganda and justify the country's indirect involvement in the conflict (Deikalo, 2023; Katerynych, 2023). Given the relatively low impact of pro-Russian information influence in Western Europe and the significant increase in its impact in Belarus, this situation poses a real security challenge, particularly in Eastern Europe.

We propose to fill this gap with a two-stage analytical approach: firstly, a network analysis based on citation patterns will uncover the interconnectedness of pro-government Telegram channels, and secondly, socio-semantic network analysis will delineate the shared concepts and discourses propagated by these channels. Our research aims to enrich and deepen our understanding of the complex interplay between pro-government and pro-Russian Telegram channels in Belarus, contributing to a more comprehensive map of Russian warfare influence.

We will answer the following research questions to achieve this goal:

RQ1. How are pro-government Telegram channels in Belarus interconnected in terms of citation patterns and core-periphery structure?

RQ2. What are the similarities and differences in the network's semantic patterns across the core, near-periphery, and far-periphery channels?

RQ3. How are the shared concepts disseminated by these channels interconnected, particularly concerning Pro-Russian and anti-Ukrainian agendas and justifications for Belarus's indirect participation in the 2022 Russian invasion of Ukraine?

The article begins by examining the relationship between Belarusian pro-government media, pro-Russian, and anti-Ukrainian agendas during the 2022 Russian invasion of Ukraine in recent academic debates. Next, we will turn to the existing studies of harmful content dissemination in social media, especially on Telegram, focusing on the promising possibilities of social-semantic network analysis. On this methodological basis, we will implement a socio-semantic network approach to examine official and grassroots pro-government Telegram channels in Belarus during the war's first and most intense month. Finally, we will turn to the patterns of mutual citation and inter-conceptual relations among the selected channels to show how pro-Russian warfare narratives were created on

the periphery among anonymous Telegram channels and further spread through odious public propagandists and several official media sources to the semantic core of the pro-government Belarusian discourse.

## 2 Literature review

### 2.1 Propaganda profile of the Belarusian pro-government media

Present state-supported propaganda in Belarus has much in common with instruments of the old Soviet regime and shares many of its achievements (Courter, 2022; Karpchuk & Yuskiv, 2021). For example, the control and management system of public opinion is based upon a monopoly on printed and audio-visual media, and repression of independent media and citizens who do not support the state's policies (Ciuriak, 2022). However, state propaganda also differs from its Soviet ancestor: it has acquired a hybrid digital-analogue nature when traditional media are complemented with cyber-propaganda (Doroshenko & Lukito, 2021).

In Belarus, social media platforms have become one of the predominant channels for distributing and receiving information. Among the plethora of social media, Telegram stands out in particular. This platform has been essential in distributing political information since 2020 and, unfortunately, has not been sufficiently researched by other authors (Kravchyk, 2022). This situation, however, changed after the 2020 protests, with an in-depth exploration of various media in Belarus, including Telegram (Wijermars & Lokot, 2022; Slobozhan et al., 2023).

Belarusian pro-government media mainly relied on newspapers and television (Silitski, 2006; Marples, 2007; Rudling, 2017; Szostek, 2018; Manaev et al., 2021). Still, after the mass protests in 2020, it turned out that the opposition movement was more successful in distributing information on social media and fully controlling the agenda (Asmolov, 2020; Laputska, 2021; Robertson, 2022; Mateo, 2022). The authorities responded by rapidly adopting unprecedented measures to restrict freedom of speech, arresting managers and rank-and-file employees of independent media, creating a register of 'extremist channels', and repressing people with interest in media that do not conform to the rhetoric of the authorities (Rust, 2022; Robertson, 2022; Onuch & Sasse, 2022). Along with the active repression, propagandists created channels on YouTube, TikTok, and Telegram to reach a broader audience and regain at least partial control over public opinion (Sulzhytski, 2022; Świerczek, 2022).

This strategy was successful. The effectiveness of spreading disinformation on social media is higher than in traditional media because it attracts a broader and more differentiated audience with regards to age and social status. In addition, information on social media is often spread virally, which is achieved, for example, with the help of bots (Ciuriak, 2022). Meanwhile, the younger generation in Belarus is less likely to trust TV sources (Alyukov, 2022). As a result, pro-government media also strive to be up to date (Doroshenko & Lukito, 2021; Martyniuk & Shuba, 2022): attracting a large number of subscribers, using slang (e.g. 'fakes'), and using new digital platforms. In addition,



pro-government channels often use various types of content, including news, analytical publications, interviews with public experts, and entertainment (Alyukov, 2022).

## 2.2 Belarusian pro-government media during the 2022 Russian invasion of Ukraine

Authors studying anti-Ukrainian narratives in Russian media emphasise that the primary audience and target of hate propaganda are Russian-speaking communities in Russia and beyond (Torichnyi et al., 2021; Kravchyk, 2022; Courter, 2022). As a result, some dubious and unverified news reaches the Western media space, especially where there are initially pro-Russian groups. At the same time, diplomats try to soften the overall picture of Russian aggression through cover-up and disinformation and convince other countries of this (Kravchyk, 2022). Western countries actively counter the spread of disinformation and propaganda through severe restrictions, such as blocking Russian TV channels, YouTube content, and other online resources (Golovchenko, 2022; Yurkova, 2018).

The situation is different with the Belarusian media: unlike their Russian counterparts, these channels are much smaller. Before 2020, Belarusian pro-government media only sometimes corresponded to the current Russian or global agenda and focused mostly on domestic and local issues. Moreover, official rhetoric towards Russia, mediated by the Belarusian pro-governmental media, was very volatile. Thus, in 2020, before the presidential election, Russian oligarchs were accused of funding the Belarusian opposition; after the August election, the agenda shifted to blaming the West for supporting the protests (Manaev et al., 2021). After the beginning of the 2020 protests, the situation changed dramatically with the arrival of media specialists from Russia to restore Lukashenka's image, which was destroyed during the elections (Wilson, 2021). In addition to frequent narratives about 'brotherly' relations with Russia, the Belarusian media field also disseminated messages framing Lukashenka as the only fighter for Belarusian independence, branding traitors from Belarusian opposition, and declaring conflict with Western countries (Sulzhitski, 2022).

Since at least 2014, Lukashenka has taken a very ambivalent position towards Ukraine (Polovyi, 2022). On the one hand, he has not expressed explicit anti-Ukrainian sentiments that have become the basis of the Kremlin's political rhetoric. However, he refrained from condemning Russia's aggressive actions. He tried to present Belarus as a peacemaker and the primary mediator of Russian-Ukrainian relations, and to benefit from cooperation with both countries. However, the 2020 protests finally turned him towards the Kremlin and its agenda (Astrouskaya, 2022; Kotljarchuk & Zakharov, 2022; Hansbury, 2023).

Although Lukashenka is still more committed to anti-Western than to anti-Ukrainian rhetoric, the media that support him are becoming more and more like their Russian counterparts (Mudrov, 2022). Since the beginning of the war, the pro-governmental agenda in Belarus tends to demonise Ukrainians, especially the Ukrainian authorities, as well as to constantly strengthen anti-Western rhetoric and create an image of Belarus as a fortress besieged by Western enemies (Katerynych, 2023; Gershovich, 2023).

Even though the pro-government media in Belarus have not been able to expand their influence significantly (Kuznetsova, 2023), they focused primarily on those who pre-

viously did not support the opposition and were in solidarity with the Lukashenka ‘geopolitical choice’ of an alliance with Russia. Thus, in the case of Belarus, the ratio of the three conditional sources of influence – local pro-Russian activists, supporters of the Lukashenka regime, and pro-Russian groups directly or indirectly subordinate to the Kremlin – remains open. We argue that in each case of the spread of aggressive pro-war narratives, there is a mix of local Belarusian supporters of war and agents of the Kremlin-facilitated agenda. While it is difficult to reliably assert the motives and goals of the degree of creative independence of these groups, we can trace the extent to which ideas, concepts and narratives are translated, modified, distorted or supplemented in the agenda of pro-Russian channels in Belarus.

### 2.3 Social and semantic network analysis of media during the 2022 war in Ukraine

The relational perspective of social network analysis offers powerful tools to map and examine complex social media structures and their role in the propagation of harmful content, including hate speech, propaganda, and disinformation (Klausen et al., 2012; Ben-David & Fernández, 2016). On the other hand, semantic analysis focuses on the meaning and context of language, providing critical insights into the subtleties of hate speech, which often relies on coded language and euphemisms (Magu & Luo, 2018; Mathew et al., 2020). Combining both approaches – social and semantic network analysis – holds the most promise (Pereira-Kohatsu et al., 2019; Nagar et al., 2023). This integrative approach enables a more comprehensive understanding of harmful content, accounting for the concepts, their context and their interplay within social networks. Researchers can develop more sophisticated and accurate models for detecting and studying harmful content in online media by focusing on social and semantic relations within hate networks.

A relational perspective on harmful content has proven to be especially effective in exploring the media and public discourse surrounding the Russia-Ukraine conflict. Several studies employed network analysis to examine the influence of different agents of information warfare on public perception. Li et al. (2023) used network analysis to investigate the social activity of the Twitter account @UAWeapons during the conflict, while Alieva et al. (2022) analysed the manipulation of political discourse by bots on social media regarding topics related to the activities of the Russian opposition. Furthermore, Ngo et al. (2022) explored public sentiment towards economic sanctions in the war using a combination of network analysis and machine learning.

Complementing these studies, papers with a stronger emphasis on textual analysis methods include semantic network analysis – Park et al. (2023), for example, use it to explore how different countries report on the same global event – combined with sentiment analysis (Tao & Peng, 2023; Eligüzel, 2023), which revealed the emotional tone of social media posts regarding the conflict.

While reinforcing the relational perspective with more sophisticated approaches, Hanley et al. (2023) used semantic search to track Russian state media narratives about the Russo-Ukrainian War on Reddit. Within the complex task of monitoring the spread of misinformation, Alieva, Ng and Carley (2022) have investigated Russian narratives about

US Biolabs in Ukraine, using network analysis to detect critical influencers and disseminators. Additionally, Džubur et al. (2022) combined sentiment and network analysis to explore Twitter discussions surrounding the conflict, identifying the most influential ‘pro-Russia’ and ‘pro-Ukraine’ hashtags and users.

Thus, most studies reviewed often use a more methodologically complex and insightful relational perspective to analyse English-language content on platforms popular in Western Europe, like Twitter. Much less attention is paid to hate communities on Telegram, one of the leading social media platforms in Russia and Belarus, where the role of propaganda and misinformation has a critical and direct impact, at least due to language and moderation features. At the same time, ‘hidden’ and local hate networks disseminating aggressive pro-war content in Belarus are almost absent from the research focus, which is mainly official pro-war propaganda.

### 3 Methodology

#### 3.1 Defining the socio-semantic network of Telegram channels

The studies mentioned above on the spread of harmful content after the 2022 Russian invasion of Ukraine reveal a split in analytical focus. This ambivalence is characterised by either an emphasis on the semantic features of the content (textual analysis) or the social relationships among online agents (network analysis). We will focus on the interplay between social and semantic networks within complex systems to bridge this divide. This text applies the socio-semantic analysis methodology (Roth & Cointet, 2010; Roth, 2013), which is appropriate for examining knowledge communities in common and exploring the blogosphere networks in particular (Roth & Cointet, 2010, p. 18; Roth, 2013, pp. 7–13).

Following the empirical protocol proposed in the socio-semantic analysis methodology (Roth & Cointet, 2010, pp. 17–18; Roth, 2013, pp. 7–8), we will conduct a combined analysis that includes both social network analysis based on observable relations among agents (such as citations and mentions), and socio-semantic network analysis, which involves the use of specific concepts (noun phrases) by agents. In our case, we define the *social network of Telegram channels* as a graph, where nodes represent Telegram channels, and links are reposts, mentions and responses to messages. Similarly, the *socio-semantic network of Telegram channels* is depicted as a bipartite graph, where one set of nodes represents channels, and the other represents concepts. The links between these two sets of nodes are expressed through the use of specific concepts by the channels. In the corresponding subsections 3.3 and 3.4, we will describe the concrete data analysis pipelines for analysing these two types of graphs in detail.

#### 3.2 Data selection procedure

Our study used the following protocol for channel selection and subsequent data extraction and processing. We used rating statistics from the TGStat service to collect statistical information about the channels, which includes a ranking section for Belarusian Telegram

channels. We then used a citation rate criterion to select the most cited channels to ensure that our sample included not only popular channels with a high number of subscribers but also less popular channels that are nevertheless often used as a source of information by other channels. In other words, our focus was on more than just audience reach, but also the channel's authority, among other channels within the pro-government network. This approach allowed us to select 21 of the most authoritative (and often popular) Belarusian official public and grassroots anonymous pro-government Telegram channels (Table 1). Specifically, our sample included seven anonymous channels, five state news agency channels, five public activist channels, and four state ministry channels. We manually selected channels with a high citation rate that directly expressed support for the government or were directly subordinated to state structures.

**Table 1** List of selected Telegram channels

Channel type	Citation index	Subscribers 22.02.2022	Subscribers 24.02.2022	Typology code
anonymous	1080	3957	5446	ANON # 1
anonymous	963	4431	4866	ANON # 2
anonymous	899	1369	1363	ANON # 3
anonymous	359	7506	8791	ANON # 4
anonymous	336	2188	2176	ANON # 5
anonymous	323	6307	6328	ANON # 6
anonymous	91	16480	14122	ANON # 7
government organisations	681	8610	9016	GOV ORG # 1
government organisations	538	24670	24676	GOV ORG # 2
government organisations	423	14428	14386	GOV ORG # 3
government organisations	410	7660	7920	GOV ORG # 4
official media	1209	22924	24,644	OFFICIAL MEDIA # 1
official media	1047	104547	113,135	OFFICIAL MEDIA # 2
official media	1004	14325	15,570	OFFICIAL MEDIA # 3
official media	842	10068	10,257	OFFICIAL MEDIA # 4
official media	635	12906	13,640	OFFICIAL MEDIA # 5
public activist	378	6463	6791	PUBLIC ACTIVIST # 1
public activist	326	7951	8054	PUBLIC ACTIVIST # 2
public activist	291	9314	9418	PUBLIC ACTIVIST # 3
public activist	115	36105	36,185	PUBLIC ACTIVIST # 4
public activist	85	20.000-21.000	20.000-21.000	PUBLIC ACTIVIST # 5

We chose data from 22 February 2022 to 24 March 2022, covering the first month of the invasion. This choice was based on several factors. Firstly, the chosen timeframe marked the initial phase of the war, a period characterised by highly fluid and viral public opinion, government positions and media narratives. Secondly, the first month of the conflict was a period of peak media activity not only in Belarus but also in Ukraine and Russia (Ghasiya & Sasahara, 2023; Kiforchuk, 2023; Ptaszek, et al., 2023). In addition, this study is one of the first, to our knowledge, to analyse the activity of pro-government Belarusian Telegram channels during the war using network analysis, so it is crucial to establish a clear and accessible baseline for further research on the dynamics of pro-government discourse in Belarus after the invasion. Finally, it was essential to choose a time frame that could be used for comparative studies, as media activity in the first month of the war has been best studied in different countries.

All messages were collected via the Telegram desktop archiving function in the .json file format without photos and audio files – the final dataset contained around 44,700 messages with text and citation information. The content of the selected channels was mainly in Russian, sometimes using Belarusian or English words. After collecting the data, we divided the initial dataset into two separate datasets – first with the citation metadata and second with messages' textual data. After constructing the final social and socio-semantic graphs, we filtered out nodes with a degree range of less than two to show only relatively significant relations.

### 3.3 Social network analysis pipeline

Our social network analysis pipeline consists of the following steps. First, we extracted only the textual part of messages and metadata from our initial dataset. Next, we discarded irrelevant data, while the citation information that was critical for our network analysis, such as 'mentions', 'forwarded\_from' and 'reply\_to\_message\_id', was processed and combined to form a 'target' column of a future network containing approximately 23,400 channel-to-channel connections. The final step was to construct a directed graph, with source nodes representing selected Telegram channels, target nodes representing cited Telegram channels, edges representing citation links, and weights based on the number of citations. The final graph was visualised and analysed using Gephi 0.10.

After creating the citation graph of 21 selected channels, we applied a mutual degree range filter in Gephi equal to 1 (one). This step allowed us to explore channels with a mutual connection to at least one other source node from our dataset. For this graph, we calculated the average degree and density metric to understand connectivity, robustness, and potential for information spread within the obtained network.

Next, we supplemented the citation graph of the selected source channels with a broader set of cited channels from mentions, responses and citations. At this stage, we used the eigenvector centrality metric to identify the most influential channels likely to be the main disseminators of information. Additionally, we manually checked the affiliation characteristics of cited/mentioned channels within the highly weighted edges in the extended citation graph.

### 3.4 Semantic network analysis pipeline

At this stage, the pipeline started with text preprocessing: removing stopwords, lemmatise words with the MyStem Python library for the Russian language and extracting meaningful concepts (noun phrases) by combining Spacy's POS tagging with the creation of bigrams and trigrams using the Gensim. Our pipeline further involved quantifying the importance of each extracted concept in our dataset using the TF-IDF algorithm. Following this, we grouped the extracted concept by channel. We then used Scikit-Learn's TfidfVectorizer to compute each channel's TF-IDF score for each concept. Next, we identified each channel's top 1000 concepts with the highest TF-IDF scores. The selected concepts were further manually checked and cleared of meaningless stop words. Lastly, we built a bipartite graph with the NetworkX library, with the channels and the top concepts serving as separate sets of nodes. Edges were drawn from each channel to its corresponding concepts, with the edge weight denoting the TF-IDF score of the concept. The final graph was visualised and analysed via Gephi 0.10.

After extracting network communities, we distinguished a particular core-periphery distribution of channels and shared concepts. We first gave a general description of channel-to-concept relations within the detected network communities. Then, we characterised the semantic core of the network and the relations between central concepts and different discourses. Finally, we indicated which concepts were on the periphery of the semantic network and analysed various hidden alliances between channels and agendas in more detail.

### 3.5 Ethical considerations

Our study navigates a delicate landscape by analysing propaganda channels on Telegram. We aim to contribute to a better understanding of these networks and their structures rather than promoting or giving visibility to these channels or their content. We are aware of the potential pitfalls associated with publicising the names of the propaganda channels, especially those that are less visible. We want to avoid inadvertently promoting or subjecting these channels to unwarranted attention. Therefore, we have kept the channel ID, names and links private while generating graphs using channel typology codes (ANON, GOV\_ORG, OFFICIAL MEDIA, PUBLIC ACTIVIST).

## 4 Findings

### 4.1 The role of Belarusian Telegram channels in the dissemination of pro-Lukashenka and Russian agendas

This section addresses the first research question (RQ1): how are pro-government Telegram channels in Belarus interconnected regarding citation patterns and core-periphery structure? We demonstrate the connectivity patterns among these channels, particularly how they cite each other and organise themselves into a structured network. We also question the extent to which these channels form a cohesive, interconnected network instead of a collection of isolated entities.



We argue that Belarusian Telegram channels act as explicit connectors in agenda-setting and intermediaries, bridging the gap between Russian media and supporters of the Lukashenka regime. This connective function can be observed at three levels: official media channels, which link the official Russian media discourse with the Belarusian government; public propaganda channels, which connect the pro-Lukashenka agenda with power groups in Russia and Belarus; and finally, anonymous activists, who are linked to the active pro-war Russian channels and further facilitate the dissemination of the pro-war agenda. Below, we will address these findings in more detail.

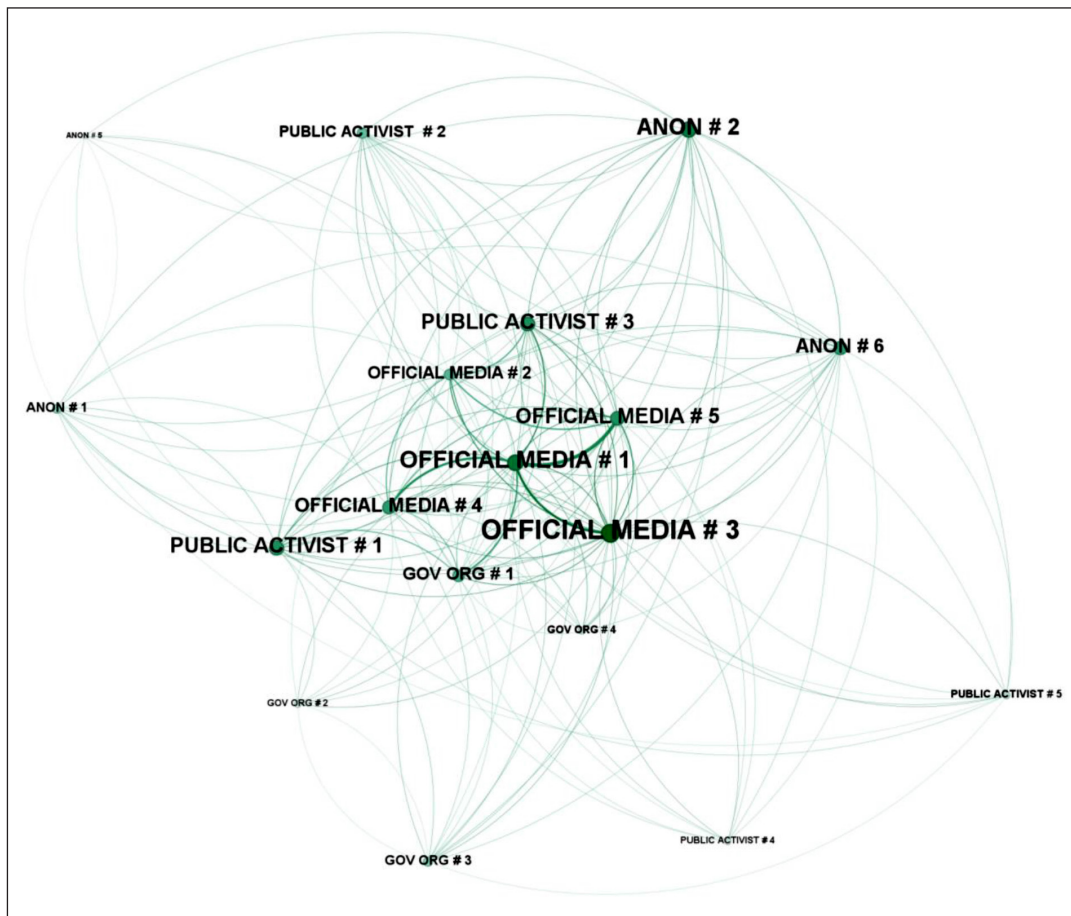


Figure 1 Co-citation network of pro-government Belarusian channels

The visualisation shows a network graph where the size of each node and the font size of the channel names indicate the centrality degree of the node. A gradient colour scheme is applied to the nodes, with dark green representing the most active channels and light green representing the least active channels. Three anonymous channels (ANON # 3, ANON # 4, ANON # 7) are missing from the graph due to lack of co-citation with other sampled channels. The central nodes of the graph are associated with official media channels. In contrast, the graph's periphery consists of nodes representing anonymous channels and channels related to less influential activists and law enforcement organisations (GOV ORG # 2, GOV ORG # 3).

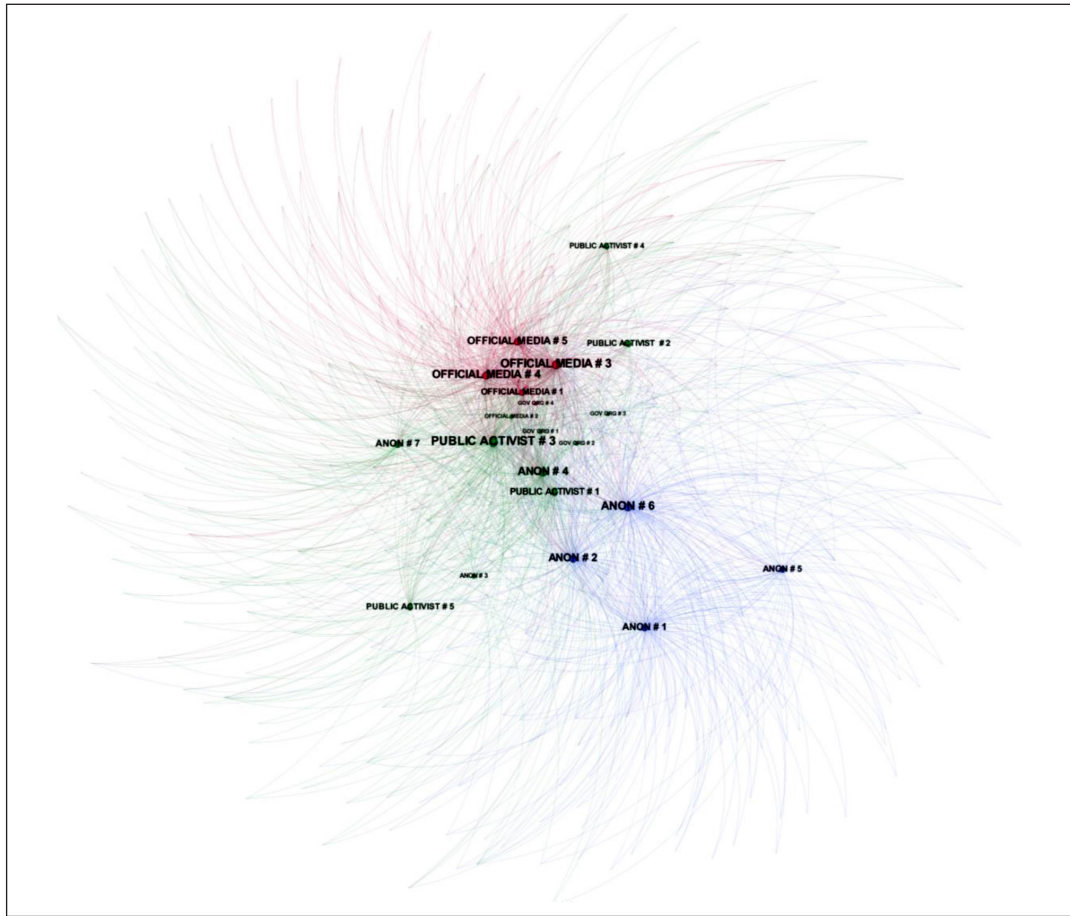


To explore connectivity patterns among the citation network, we examined a mutual citation graph containing 18 out of 21 interconnected source nodes, with an average degree of 8.5 and a density of 0.5. These characteristics indicate that the selected source channels form a highly interconnected, dense network, uniting official pro-government channels with public activists and anonymous channels. The core of the citation network is formed by four official news media channels that serve as the primary source of information for the anonymous channels and government agency channels located on the periphery (Figure 1).

However, by expanding the initial graph of 21 selected channels with a broader set of citations and mentions, it is possible to see the implicit role of Russian media as the peripheral source of information dissemination (Figure 2). Based on the eigenvector centrality metric examination, the top ten most central channels include seven Belarusian channels and three directly related to major Russian media. Moreover, eight of the ten most-weighted edges have nodes representing links between Belarusian news media and the leading Russian state news media. This trend continues, including high-weighted connections with numerous unofficial, anonymous Russian channels promoting an overt pro-war agenda.

Altogether, it is possible to identify three network communities within which information spreads in the citation network. The *pro-Lukashenka community* shows a strong alignment with the Belarusian security forces and includes channels that engage in hate speech against the opposition. Notably, the central channels of this community also include those associated with Russian warfare correspondents, politicians, and Putin's administration. The *grassroots pro-Russian community* is characterised by its anonymous nature and support for Russian military efforts in Ukraine, with a significant presence of pro-war content. This community has few Belarusian channels and poor connections with official Belarusian media or departments. Finally, the *Belarusian state media community* presents a mix of Belarusian and Russian state media sources, indicating a collaborative attitude towards the relations between Russia and Belarus, with a more diversified approach in its citations and connections. This community is mainly formed by official agents associated with various pro-government organisations, civil initiatives, opinion leaders and politicians in Belarus and Russia.

Finally, according to the core-periphery structure, the core's most authoritative nodes primarily include channels associated with state-controlled Belarusian media, pro-Lukashenka public activists, and security organisations. The periphery, in contrast, consists of grassroots pro-Russian channels. These agents are likely not directly subordinate to the Lukashenka regime or its power departments, even though they express loyalty on a discursive level. Moreover, the periphery emerges as the most active pro-war and anti-Ukrainian pro-Russian content disseminator, though still maintaining citation autonomy even from Russian official media.



**Figure 2** Network Communities in the Citation Network

This network graph shows the centrality and connectedness of different media channels, with the size of the nodes and font size indicating the harmonic closeness centrality of each channel. The largest nodes represent the most connected selected channels within each network community. The graph is colour-coded to identify three distinct communities:

Green, 40 percent of all channels: Belarusian pro-Lukashenka community, including government media, loyal activists and law enforcement agencies, which also have links to Russian political figures and media. Overall, the ratio of Belarusian to Russian channels is roughly equal, with a dominance in the core of the Belarusian security forces and public activists.

Blue, 35 percent of all channels: Grassroots pro-Russian community, claiming Belarusian affiliation but mostly citing Russian sources. Mostly anonymous, with frequent links to Russian military experts, war correspondents and media from the Donetsk and Luhansk People's Republics. The agenda of the quoted channels openly supports the war in Ukraine.

Red, 25 percent of all channels: Belarusian state media community, with a balanced mix of Belarusian and Russian state news media as cited sources. In general, the source base for the official pro-government Belarusian channels here is more diversified, representing official news media and channels associated with registered organisations and subordinate civil initiatives, opinion leaders and politicians in Belarus and Russia.

## 4.2 Peripheral integration of pro-war and anti-Ukrainian agendas within Belarusian pro-government channels

### *Inter-conceptual relations between pro-government channels*

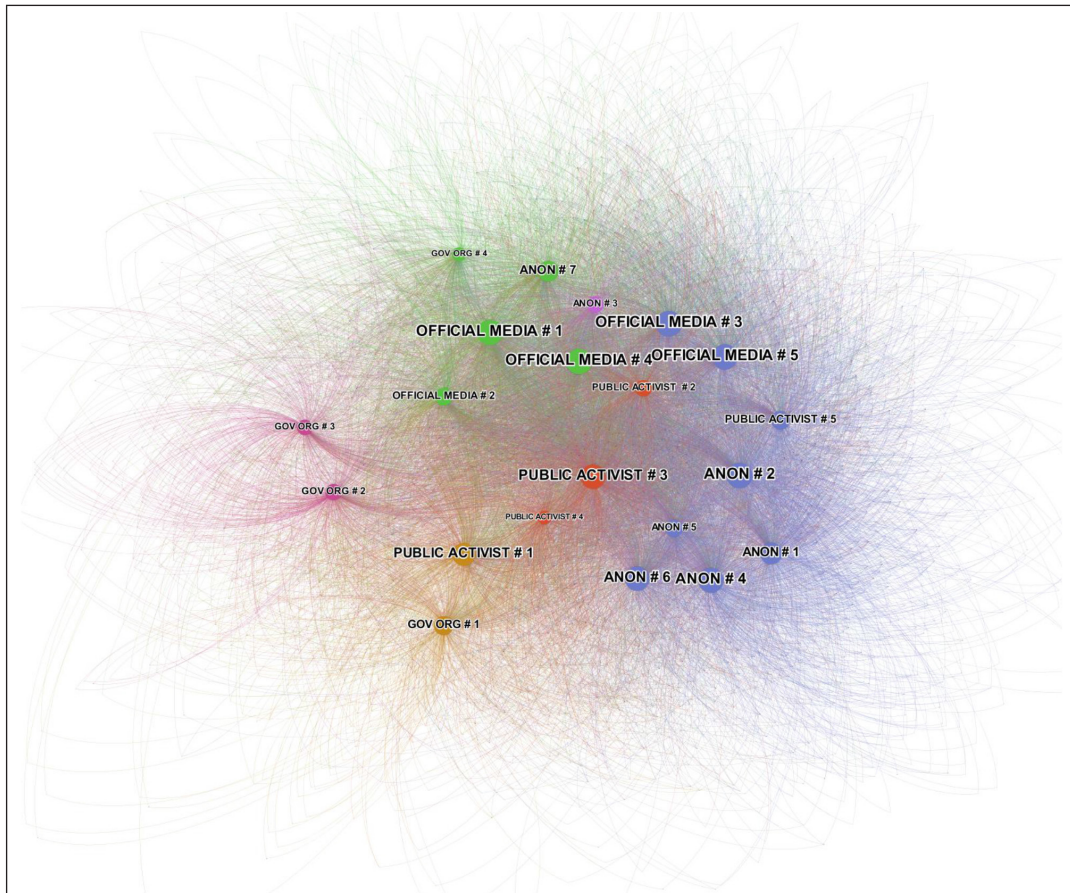
In this section, we will answer RQ2: what are the similarities and differences in the network's semantic patterns across the core, near-periphery, and far-periphery channels? Employing a bipartite graph analysis that links the Telegram channels to their dominant keywords, we first identify the most common shared concepts across the network. We then highlight the available network communities based on how channels use certain shared concepts. Finally, we examine keyword usage patterns at the core and periphery levels of the network. Thus, the main focus of this section is to trace the dissemination of pro-war and anti-Ukrainian concepts within the semantic field of pro-government Belarusian Telegram channels.

We argue that there is a group of semantically related Telegram channels within the structure of the Belarusian pro-government network, which is mainly focused on spreading anti-Ukrainian and pro-war concepts from the Russian media discourse. This is expressed in a dual-core structure in the semantic network of Lukashenka supporters, characterised by the tension between pro-Russian military and pro-Belarusian state-centred poles. However, deeper inter-conceptual relations and the diffusion of pro-war and anti-Ukrainian agendas into pro-government discourse can be seen not in the centre, where the official and state-centric ones balance pro-Russian/pro-war channels, but in the periphery, where the shared concepts denote relatively more isolated and aggressive ideas rather than common themes. Next, we examine this conclusion in more detail.

The previous community structure showed a transparent state-centric core and a pro-Russian grassroots periphery, considering only the observable social links between channels, such as citations and mentions. However, this structure becomes more diverse when focusing on the mutual use of terms, as shown in Figure 3. First, the core-periphery relationship is changing: due to the specific agenda and discursive isolation, channels of official state agencies, including those of police and military organisations, are shifting towards the far periphery. Secondly, as narratives related to the war in Ukraine spread rapidly after the invasion, pro-Russian grassroots channels, which were located on the periphery of *the social network*, moved towards the core of the *socio-semantic network*. In addition, two official media channels demonstrated intensive ties with the pro-Russian community due to their frequent use of war-related concepts predominantly framed by Russian media discourse.

These shifts suggest that the pro-Russian agenda is beginning to dominate within the network structure, leading to a fragmentation of the previously state-centric core, as shown in Figure 4. Specifically, the most extensive community, focused on pro-Russian warfare, focuses on keywords related to Ukraine, Russia, the US and warfare. Accompanying this central theme are significant peripheral concepts that resonate with more aggressive pro-war propaganda: 'Nazi', 'Banderovite', 'Biolab', 'militant', 'special military operation', 'Odesa' and 'Kherson'. This is followed by the pro-government community, with a prevalence of official media channels. This community is closely linked to pro-Lukashenka political and ideological terms at its core – 'Lukashenka', 'Belarus', 'citizen' – alongside

peripheral terms related to domestic political media keywords, such as 'COVID-19', 'economy', 'foreign ministry', 'constitution', 'minister' and 'chairman'. The law enforcement community focuses on operational and security-related terms, reflecting a strong interest in internal security and the societal impact of law enforcement. The community of public pro-Lukashenka activists uses more confrontational grassroots rhetoric in support of Lukashenka, framing his conflict with the West and suppression of the opposition in terms of 'world', 'side', 'political', 'West', 'state' and 'people', along with 'fighter' / 'zmagar'. Finally, the Belarusian military community focuses on national defence and army engagements, reflecting a devotion to national security and strategic military concerns.



**Figure 3** Channel-concept network communities in the bipartite semantic network

The graph represents different network communities in the socio-semantic network, linking Telegram channels with shared concepts:

- Blue (28.92%): the pro-Russian war community, centred on the 2022 Russian invasion with the prevalence of nodes with a strong pro-war focus.
- Green (20.98%): the pro-government community, with a focus on Belarus-centric and Lukashenka-aligned narratives, extending to domestic issues.
- Red (16.35%): Indicates the public pro-Lukashenka activists, characterised by assertive language against Western opposition and in support of the current regime.





- Green, the pro-government community: Belarus, president, Belarussian, negotiations, Lukashenka, citizen, the First (Lukashenka), work, Ministry of External Affairs, referendum, Minsk, covid\_19, RB (Republic of Belarus), constitution, (voting) observer, organisation, chairman, development, etc.
- Red, public pro-Lukashenka activists: world, West, Donbass, conflict, child, western, Ukrainian (citizen), Nazi, the people, state, life, political, Belarussian (citizen), political scientist, denazification, fighter (zmagar, jargon), Kyiv regime, etc.
- Purple, the Law Enforcement Community: criminal, (police) staff member, chief, action, road accident, investigation, militia, investigator, operative agent, psychotropic, drug, criminal code, crime, suspect, unlawful, drug control, murder, defendant in a criminal case, etc.
- Yellow, The Belarussian military community: military, force, Republic of Belarus, military personnel, subdivision, brigade, photo, Vayar Military News Agency, combat, task, Independent Guards Regiment, (military) service, VoenTV (Military TV), border crossing point, major general, General Staff, state border, ideological work, defence minister, officer, polygon, permanent dislocation, etc.
- Pink: A separate anonymous war-oriented channel (ANON # 3): GUBOPiK, O/in 'on Ukraine' vs 'in Ukraine', fuck, truth, activity, (Konstanty) Kalinowski, Turkey, internet, Twitter, blakitinyy (blue, jargon), Pole, Intelligence, etc.

#### *The core-periphery structure of the network*

As seen above, the use of concepts by different channels not only has a visible community structure, but also demonstrates the distance between the more homogeneous and semantically neutral core concepts and the more diverse (as well as aggressive) peripheral concepts. Furthermore, the network shows the existence of the dual-core structure, indicating a peculiar tension between the two separate discourses articulated by various Lukashenka supporters, as illustrated in Table 2. At the first pole, there is a semantic fusion between Lukashenka supporters and pro-Russian military narratives. In contrast to the first pole, which is aligned with Russia against the West, the second pole emphasises an independent Belarussian state with Lukashenka as its institutional foundation, focusing on domestic stability and governance. This discursive alliance becomes even more possible with the channels of public propagandists serving as the primary mediators between the internal pro-state Belarussian discourse and the external pro-Russian military discourse.

But even more connections between the pro-Russian and pro-Lukashenka channels are made on the periphery through various shared concepts that indicate more specific, diverse, and often aggressive ideas, themes and agendas. For example, one of the semantic alliances was built around the idea of fighting 'nationalists' in Ukraine and Belarus. Here, pro-Lukashenka and pro-Russian channels used two sets of concepts from the Russian 'special operations' vocabulary: official Russian military language ('Ukrainian nationalist', 'Ministry of Defence briefing', 'foreign mercenary', 'strategic' and 'missile strike') was combined with anti-Ukrainian propaganda terms such as 'Maidan', 'propaganda', 'Kyiv regime', 'denazification', 'Crimea', 'SBU (Security Service of Ukraine)' and 'victory'.

**Table 2** Two poles of socio-semantic network discourse articulated by Lukashenka supporters

Discursive affiliation	Shared concepts
pro-Russian / pro-Lukashenka core	'Putin', 'Russian', 'sanctions', 'negotiations', 'Moscow', 'war', 'West', 'conflict', 'Donbas', 'Ukrainian'.
Pro-government / pro-Lukashenka core	'Belarusians', 'state', 'president', 'civilian', 'work', 'action', 'protection', 'Belarusian', 'Belarus', and 'management'.

### 4.3 The relational model of Belarusian pro-government discourse

To answer RQ3, we created a relational discourse model grounded on a unipartite projection of a bipartite semantic network of inter-conceptual relations among different channels. This model represents the idea of semantic similarity between distinct concepts based on their co-occurrence across various channels. Notably, this model clarifies possible relational patterns of content creation and dissemination, rather than explicitly describing more complex pro-government discourse.

In this section, we conclude that the Belarusian pro-government Telegram network not only disseminates state-centred narratives but is also heavily influenced by pro-Russian media discourse at several levels. While there is a distinct space for domestic political discourse that is less influenced by Russian discourse, a strong emphasis on supporting Russia's military and foreign policy agendas remains dominant. However, while there is a clear tendency to import military and anti-Ukrainian narratives through anonymous channels and channels of public propagandists, there is also some selective engagement with pro-Russian content.

We seek to clarify these findings by exploring four network communities that represent separate but interconnected semantic fields indicative of the heterogeneous interplay of agendas, ideas, and concepts within the pro-government discourse landscape, contingent on the dissemination and sharing of these concepts across channels:

1. Pro-Russian Warfare field (30.87 percent of concepts): include 'Ukraine', 'Russia', 'USA', 'military', 'Kyiv', 'Zelenskyy', 'NATO', 'Donbas', 'fake', 'power', 'child', 'Mariupol', 'weapon', 'operation', 'Poland', 'district', 'AFU' (Armed Forces of Ukraine). This part of the semantic network remains the main source of anti-Ukrainian propaganda in Belarus and is used to represent Russia's supporters in this war. These concepts are mostly connected with pro-Russian anonymous channels in Belarus and with channels in Russia, which are used as a source of quotations for content creation.
2. Pro-Russian Foreign Policy field (28.7 percent of concepts): include 'Russia', 'territory', 'negotiations', 'RF' (Russian Federation), 'Putin', 'situation', 'decision', 'sanctions', 'EU', 'Europe', 'head', 'meeting', 'representative'. In this community, we can see the process of merging the Belarusian state discourse with the foreign policy agenda of the Kremlin, which is disseminated by some official Belarusian channels through quotations from Russian media. Thus, the semantics of Russian foreign policy completely replace the language of Belarusian international relations.



3. Belarusian Domestic Policy (28.1 percent of concepts): includes 'Belarus', 'force', 'today', 'citizen', 'first', 'Republic of Belarus', 'relationship', 'work', 'military serviceman', 'action', 'resident', 'connection', 'employee', 'Minsk', 'security'. The language used to describe internal politics is less affected by the pro-Russian agenda. It deals mainly with topics related to the activities of the government and other official organisations. This community is more connected to the official channels of the pro-government pole of the semantic core, which rarely use content from pro-Russian Telegram channels or address the topic of the war in Ukraine.
4. Pro-Lukashenka field (12.26 percent of concepts): key concepts include 'president', 'Belarusian', 'Lukashenka', 'side', 'war', 'peace', 'border', 'west', 'against', 'conflict', 'western', 'people', 'state', 'special operation', 'Belarusian', 'political'. Here we see a fusion of aggressive pro-Lukashenka rhetoric with pronounced anti-Western and pro-Russian agendas. The primary sources of these concepts are the channels connected with Lukashenka's press service, which show the position of Lukashenka himself: Belarus is drawn into the conflict and wants to preserve itself in the global war between the West and Russia but, at the same time, remains an ally of Russia.

Thus, we can identify an overall semantic pattern that emphasises a strong connection between Belarusian state discourse and the Russian media field, together with an attempt to maintain the visibility of the national narratives articulated by the pro-Lukashenka field. Observed fragmentation may point to the difficulty, or even impossibility, of keeping coherent narratives of national identity or state legitimacy separate from the overarching influence of Russian discourse in Belarus.

## 5 Discussion

The findings of this study highlight the ambivalent role of pro-government Telegram channels in Belarus in disseminating narratives supportive of both the Lukashenka regime and Russia's foreign policy, particularly during the 2022 invasion of Ukraine. By using a combined approach of network and semantic analysis, this research reveals an interplay between local Belarusian and external Russian media agents, highlighting how pro-Russian Telegram channels in Belarus serve not only as platforms for the dissemination of state-aligned messages, but also as active propagators of war-supporting statements.

Crucially, our analysis identifies a dual-core network structure within the discourse, highlighting a tension between pro-Russian military narratives and those emphasising Belarusian state interests. This structure reflects a more complex media landscape, where pro-government channels in Belarus offer a mixed agenda between national sovereignty discourse and alignment with Russian geopolitical influence. The significant peripheral presence of Russian channels promoting pro-war content further complicates this dynamic, reflecting a tendency to spread support for military action against Ukraine and shape public opinion in favour of the Russo-Belarusian alliance within the core of pro-government media discourse.

Separately, we would like to compare our proposed relational model of Belarusian pro-government discourse with the findings published in Petro Katerynych's recent paper "Propaganda at Play: A Thematic Analysis of Belarusian Media Narratives in the Context of the Russo-Ukrainian War" (Katerynych, 2023). The thematic analysis in that paper complements our relational model by thoroughly examining the content and themes within Belarusian state media, highlighting the explicit promotion of narratives that support Russian foreign policy, demonise Ukraine, and criticise Western intervention. Considering the findings of both papers, we argue that the accession of Belarusian media to Russian military narratives complements concerns about the autonomy of Belarusian public discourse and Russia's increasing influence on domestic and foreign policy decision-making. While such concerns have been raised repeatedly (Hansbury, 2023; Katerynych, 2023; Kotljarchuk & Zakharov, 2022), our study can serve as an empirical illustration of how the vast periphery of Russian Telegram channels influences Belarusian pro-government media discourse at a structural level.

However, despite the potential of the approach presented, we need to be aware of several methodological pitfalls. First, our time frame is limited to the first month of the war. It cannot reveal the long-term evolution of the semantic and social relations between the channels, as the discourse will likely change throughout the conflict. Second, the paper focused only on prominent and easily identifiable pro-government channels, possibly excluding less-visible channels or those blocked by Telegram. Thirdly, the methodological approach prioritised citation patterns within the social network, which inherently omits other possible channel relationships. In addition, the semantic network analysis only included significant noun phrases, thus simplifying the potential complexity of semantic relationships. Finally, the potential biases introduced by community detection on bipartite graphs or the transformation from a bipartite to a unipartite projection, which inevitably leads to a loss of information and simplifies the complex relationships between channels and concepts, must be acknowledged.

Future research should explore these narratives' impact on public opinion within Belarus and the wider region, examining how the dissemination of pro-government and pro-Russian content influences attitudes towards Ukraine, Russia, and Belarusian identity. An additional focus on the counter-narratives presented by opposition channels and their role in the information ecosystem could also provide valuable insights into the resilience of civil society discourse in Belarus.

## 6 Conclusion

The dominance in the academic literature of the image of monolithic and unified state propaganda in Belarus, formed by cooperation between official media or anonymous channels, has left the factor of local (and possibly independent of Minsk) pro-Russian agents in the shadows. These agents create narratives by rethinking the Kremlin's propaganda stamps and mixing them with Belarusian ideological and cultural resources. As a result of underestimating the various peripheral agents involved in information dissemination within the pro-government network in Belarus, hidden grassroots and anonymous

information sources remain to be explored. Given that many of these agents create unique ways of representing and disseminating group hatred and providing ideological support for the war in Ukraine, underestimating this factor poses a severe threat to regional security in Eastern Europe. This paper contributes to a more nuanced understanding of communication strategies in Eastern Europe by illustrating how the network of pro-government Telegram channels in Belarus does not operate as an isolated and monolithic state-controlled structure but as part of a broader, highly interconnected system of information dissemination between Russian and Belarusian media agents.

It is crucial to emphasise that pro-government Telegram channels in Belarus appeared as a complex social network of interconnected nodes, characterised by distinct citation patterns and a defined core-periphery structure. This network not only facilitates the dissemination of the pro-government agenda but also serves as a bridge between pro-Russian peripheral agents and supporters of the Lukashenka regime. At the level of mutual citations, we can observe a dense and cohesive network, mainly concentrated around official news media channels, which form the core and are linked to both pro-government public activists and anonymous channels. However, extending the citation network with a broader set of external citations shows that the periphery, characterised by grassroots pro-Russian channels and less connected to the official Belarusian media, plays a crucial role in spreading pro-war and anti-Ukrainian sentiments. Thus, besides the apparent use of Telegram channels by the Belarusian government and its allies to promote a unified pro-Lukashenka agenda, we can observe the significant peripheral influence of Russian Telegram channels within this network.

Expanding the explicit citation analysis with the implicit word usage patterns highlights the centrality of the Ukrainian conflict in the network's discourse. Moreover, there is varied evidence of the dissemination of explicit anti-Ukrainian hate propaganda on different levels of the network structure. The presence of discrete communities within the network, each with a specific thematic focus, adds complexity to the observed semantic landscape. In addition to semantically diverse groups, a dual-core structure alternates between pro-Russian military narratives and pro-Belarusian state-centric narratives. Although a core balance exists between state-centric channels and those aligned with the pro-Russian military agenda, a more dynamic interchange of concepts is observed in the periphery. The given activity suggests the possibility of disseminating pro-Russian military discourse within the Belarusian network, which is state-centric, through anonymous sources that are broader and independent of the government.

The study effectively uncovers the main consequences of such a structural integration with the plethora of peripheral pro-Russian channels by applying a relational discourse model based on a unipartite projection of a bipartite semantic network. In other words, the research reveals a critical dominance of pro-Russian media discourse in Belarusian pro-government channels, which is evident at multiple levels and across various semantic fields. While there is still an area for domestic political discourse that is less influenced by Russian narratives, the overall trend is towards spreading pro-war / anti-Ukrainian agendas and using Russian foreign policy terminology. The analysis of four semantic fields – Pro-Russian Warfare, Pro-Russian Foreign Policy, Belarusian Domestic Policy, and Pro-Lukashenka – indicates that almost 60 percent of the meaningful shared concepts belong to the Russian political discourse.

The idea of monolithic, centralised, Kremlin- or Minsk-supervised propaganda in Belarus is somewhat misleading. Although there is an anti-Western consensus among both public supporters of the regime and anonymous Telegram channels, this rhetoric is mixed with both pro-Russian and anti-Ukrainian agendas to varying degrees. Official state media are less likely to express explicitly anti-Ukrainian views, while anonymous Telegram channels are much more active in spreading group hatred. After the beginning of the Russian invasion, hostility towards Ukraine and an openly pro-Kremlin position remain markers that allow for an at least analytical distinction between grassroots pro-Russian, Belarusian pro-government, and Kremlin-supported discourse. In the case of Belarus, the primary sources of the creation and dissemination of group hatred towards Ukrainians are grassroots, anonymous, and public pro-government activists who are strongly connected to the pro-Russian channels.

## Acknowledgements

The work on this paper was financed by the Austrian Academy of Sciences' go!digital 3.0 program.

## References

- Alieva, I., Moffitt, J. D., & Carley, K. M. (2022). How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. *Social Network Analysis and Mining*, 12(1), 80. <https://doi.org/10.1007/s13278-022-00908-6>
- Alieva, I., Ng, L. H. X., & Carley, K. M. (2022, December). Investigating the Spread of Russian Disinformation about Biolabs in Ukraine on Twitter Using Social Network Analysis. In *2022 IEEE International Conference on Big Data (Big Data)* (pp. 1770–1775). IEEE. <https://doi.org/10.1109/BigData55660.2022.10020223>
- Alyukov, M. (2022). Propaganda, authoritarianism and Russia's invasion of Ukraine. *Nature Human Behaviour*, 6(6), 763–765. <https://doi.org/10.1038/s41562-022-01375-x>
- Asmolov, G. (2020). The path to the square: the role of digital technologies in Belarus' protests. *Open Democracy*. <https://www.opendemocracy.net/en/odr/path-to-square-digital-technology-belarus-protest/>
- Astrouskaya, T. (2022). In Schoolbooks and on Telegram: What Is the Place of Ukraine and Ukrainians in the Memory Politics of Post-Soviet Belarus?. *Journal of Applied History*, 4(1–2), 9–27. <https://doi.org/10.1163/25895893-bja10033>
- Ben-David, A., & Matamoros-Fernández, A. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, 10, 1167–1193.
- Chen, E., & Ferrara, E. (2023, June). Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between Ukraine and Russia. In *Proceedings of the International AAAI Conference on Web and Social Media*, 17(1), 1006–1013. <https://doi.org/10.1609/icwsm.v17i1.22208>

- Ciuriak, D. (2022). The role of social media in Russia's war on Ukraine. *Available at SSRN 4078863*. <http://dx.doi.org/10.2139/ssrn.4078863>
- Courter, I. J. (2022). Russian Preinvasion Influence Activities in the War with Ukraine. *Military Review*, 1, 16–27. <https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/September-October-2022/Courter/> Accessed: 27.06.2023.
- Deikalo, E. (2023). BHR Agenda and Authoritarian Regimes: The Case of Political and Human Rights Crisis in Belarus Since 2020. *Business and Human Rights Journal*, 9(1), 150–156. <https://doi.org/10.1017/bhj.2023.8>
- Doroshenko, L., & Lukito, J. (2021). Trollfare: Russia's disinformation campaign during military conflict in Ukraine. *International Journal of Communication*, 15, 28. 4662–4689.
- Džubur, B., Trojer, Ž., & Zrimšek, U. (2022). Semantic Analysis of Russo-Ukrainian War Tweet Networks. *SCORES: Ljubljana, Slovenia*.
- Eligüzül, İ. M. (2023). Russia-Ukraine Conflict: A Text Mining Approach through Twitter. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 12(1), 272–291. <https://doi.org/10.17798/bitlisfen.1238241>
- Geissler, D., Bär, D., Pröllochs, N., & Feuerriegel, S. (2023). Russian propaganda on social media during the 2022 invasion of Ukraine. *EPJ Data Science*, 12(1), 35. <https://doi.org/10.48550/arXiv.2211.04154>
- Gershovich, L. (2023). Anti-Semitism in the Propaganda and Public Discourse in Russia, Ukraine, and Belarus during the Russia-Ukraine War (February–August 2022)–ACTA. *Analysis of Current Trends in Antisemitism-ACTA*, 42(1), 1–43. <https://doi.org/10.1515/actap-2023-2001>
- Ghasiya, P., & Sasahara, K. (2023). Messaging strategies of Ukraine and Russia on Telegram during the 2022 Russian invasion of Ukraine. *First Monday*, 28(8), 1–24. <https://doi.org/10.5210/fm.v28i8.12873>
- Golovchenko, Y. (2022). Fighting propaganda with censorship: A study of the Ukrainian ban on Russian social media. *The Journal of Politics*, 84(2), 639–654. <https://doi.org/10.1086/716949>
- Greene, S. A. (2022). You are what you read: media, identity, and community in the 2020 Belarusian uprising. *Post-Soviet Affairs*, 38(1–2), 88–106. <https://doi.org/10.1080/1060586X.2022.2031843>
- Greene, T. (2023). Natural allies? Varieties of attitudes towards the United States and Russia within the French and German radical right. *Nations and Nationalism*. <https://doi.org/10.1111/nana.12957>
- Hanley, H. W., Kumar, D., & Durumeric, Z. (2023, June). Happenstance: Utilizing Semantic Search to Track Russian State Media Narratives about the Russo-Ukrainian War On Reddit. *Proceedings of the International AAAI Conference on Web and Social Media*, 17(1), 327–338. <https://doi.org/10.1609/icwsm.v17i1.22149>
- Hansbury, P. (2023). *Belarus in Crisis: From Domestic Unrest to the Russia-Ukraine War*. Oxford University Press. <https://doi.org/10.1093/oso/9780197747704.001.0001>
- Jarynowski, A. J. (2022). Conflicts driven pandemic and war issues in social media via multi-layer approach of German twitter. *Interdisciplinary Research*, 1–9. <https://doi.org/10.1145/3583780.3615468>

- Karpchuk, N., & Yuskiv, B. (2021). Dominating Concepts of Russian Federation Propaganda Against Ukraine (Content and Collocation Analyses of Russia Today). *Politologija*, 102(2), 116–152. <https://doi.org/10.15388/Polit.2021.102.4>
- Katerynych, P. (2023). Propaganda at Play: A Thematic Analysis of Belarusian Media Narratives in the Context of the Russo-Ukrainian War. *Media Literacy and Academic Research*, 6(1), 23–39. <https://doi.org/10.34135/mlar-23-01-02>
- Kiforchuk, K. (2023). Frequency Analysis of Russian Propaganda Telegram Channels. *Theoretical and Applied Cybersecurity*, 5(1), 80–86. <https://doi.org/10.20535/tacs.2664-29132023.1.277647>
- Kizlova, K., & Norris, P. (2022). What do ordinary Russians really think about the war in Ukraine?. *LSE European Politics and Policy (EUROPP) blog*. <https://blogs.lse.ac.uk/europpblog/2022/03/17/what-do-ordinary-russians-really-think-about-the-war-in-ukraine/> Accessed: 01-07-2023.
- Klausen, J., Barbieri, E. T., Reichlin-Melnick, A., & Zelin, A. Y. (2012). The YouTube Jihadists: A social network analysis of Al-Muhajiroun's propaganda campaign. *Perspectives on Terrorism*, 6(1), 36–53. <http://www.jstor.org/stable/26298554>
- Kotljarchuk, A., & Zakharov, N. (2022). Belarus' relations with Ukraine and the 2022 Russian invasion: Historical ties, society, and realpolitik. *Baltic Worlds*, (1-2), 32-37.
- Kravchyk, V. (2022). How Russian Diplomatic Missions Spread Propaganda. *Ukraine Analytica*, 1(27), 44–51.
- Kuznetsova, D. (2023). Broadcasting Messages via Telegram: Pro-Government Social Media Control During the 2020 Protests in Belarus and 2022 Anti-War Protests in Russia. *Political Communication*, 41(4), 509–530. <https://doi.org/10.1080/10584609.2023.2233444>
- Laputska, V. (2021). *The power of internet as a game-changer for Belarusian protests*. *New Eastern Europe*, 5(48), 42–47.
- Lawriwsky, M. (2023). The worldwide success of Russian propaganda. *Quadrant*, 67(6), 53-59. <https://search.informit.org/doi/10.3316/informit.060527169864075>
- Li, Q., Liu, Q., Liu, S., Di, X., Chen, S., & Zhang, H. (2023). Influence of social bots in information warfare: A case study on @UAWeapons Twitter account in the context of Russia–Ukraine conflict. *Communication and the Public*, 8(2), 54–80. <https://doi.org/10.1177/20570473231166157>
- Magu, R., & Luo, J. (2018, October). Determining code words in euphemistic hate speech using word embedding networks. *Proceedings of the 2nd workshop on abusive language online (ALW2)*, 2018(10), 93–100. <https://doi.org/10.18653/v1/W18-5112>
- Manaeu, O., Rice, N., & Taylor, M. (2021). The evolution and influence of Russian and Belarusian propaganda during the Belarus presidential election and ensuing protests in 2020. *Canadian Slavonic Papers*, 63(3–4), 371–402. <https://doi.org/10.1080/00085006.2021.1997285>
- Marples, D. R. (2007). *The Lukashenka phenomenon-Elections, Propaganda, and the Foundations of Political Authority in Belarus*. Trondheim Studies on East European Cultures and Societies. No. 21. [https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2711454/Nr%2021\\_B5.pdf](https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2711454/Nr%2021_B5.pdf)
- Martyniuk, O., & Shuba, A. (2022). The Official State Narrative on the Belarusian Protests of 2020 and Its Correlations to Non-Official Uncredited Belarusian Media with Foreign Representations. *European Journal of Transformation Studies*, 10(2), 89–102. <https://czasopisma.bg.ug.edu.pl/index.php/journal-transformation/article/view/8212>



- Mateo, E. (2022). "All of Belarus has come out onto the streets": exploring nationwide protest and the role of pre-existing social networks. *Post-Soviet Affairs*, 38(1–2), 26–42. <https://doi.org/10.1080/1060586X.2022.2026127>
- Mathew, B., Illendula, A., Saha, P., Sarkar, S., Goyal, P., & Mukherjee, A. (2020). Hate begets hate: A temporal study of hate speech. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2), 1–24. <https://doi.org/10.1145/3415163>
- Mejias, U. A., & Vokuev, N. E. (2017). Disinformation and the media: the case of Russia and Ukraine. *Media, culture & society*, 39(7), 1027–1042. <https://doi.org/10.1177/0163443716686672>
- Metzger, M. M. (2023). Authoritarian media abroad: the case of Russia and RT News. In N. Seltzer & S. Wilson (Eds.) *Handbook on Democracy and Security* (pp. 126–140). Edward Elgar Publishing. <https://doi.org/10.4337/9781839100208.00015>
- Mudrov, S. A. (2022). "We did not unleash this war. Our conscience is clear". The Russia–Ukraine military conflict and its perception in Belarus. *Journal of Contemporary Central and Eastern Europe*, 30(2), 273–284. <https://doi.org/10.1080/25739638.2022.2089390>
- Nagar, S., Barbhuiya, F. A., & Dey, K. (2023). Towards more robust hate speech detection: using social context and user data. *Social Network Analysis and Mining*, 13(1), 47. <https://doi.org/10.1007/s13278-023-01051-6>
- Navumau, V. (2020). Integration or absorption? Analyzing the propagandist narratives generated by Russia-backed online regional media in Belarus. *Demokratizatsiya: The Journal of Post-Soviet Democratization*, 28(3), 461–484. <https://muse.jhu.edu/article/762317>
- Nazaruk, T. (2022). Subscribe and follow. telegram and responsive archiving the war in ukraine. *Sociologica*, 16(2), 217–226. <https://doi.org/10.6092/issn.1971-8853/15339>
- Ngo, V. M., Huynh, T. L., Nguyen, P. V., & Nguyen, H. H. (2022). Public sentiment towards economic sanctions in the Russia–Ukraine war. *Scottish Journal of Political Economy*, 69(5), 564–573. <https://doi.org/10.1111/sjpe.12331>
- Oates, S., Lee, D., & Knickerbocker, D. (2022). Data Analysis of Russian Disinformation Supply Chains: Finding Propaganda in the US Media Ecosystem in Real Time. Oates, Sarah, Doowan Lee, and David Knickerbocker. <https://dx.doi.org/10.2139/ssrn.4218316>
- Onuch, O., & Sasse, G. (2022). The Belarus crisis: people, protest, and political dispositions. *Post-Soviet Affairs*, 38(1–2), 1–8. <https://doi.org/10.1080/1060586X.2022.2042138>
- Park, H., Pak, J., & Kim, Y. (2023). Analysis of News Article Various Countries on a Specific Event Using Semantic Network Analysis. In *International Conference on Information Technology-New Generations* (pp. 229–235). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-031-28332-1\\_26](https://doi.org/10.1007/978-3-031-28332-1_26)
- Pereira-Kohatsu, J. C., Quijano-Sánchez, L., Liberatore, F., & Camacho-Collados, M. (2019). Detecting and monitoring hate speech in Twitter. *Sensors*, 19(21), 4654. <https://doi.org/10.3390/s19214654>
- Polovyi, T. (2021). Deterioration of Polish-Belarusian Relations after the 2020 Political Crisis in Belarus. *Ukrainian Policymaker*, 9(9), 93–100. <https://doi.org/10.29202/up/9/9>
- Polovyi, T. Y. (2022). The Republic of Belarus in Russian-Ukrainian War in 2014–2022. *Publishing House "Baltija Publishing"*. <https://doi.org/10.30525/978-9934-26-223-4-128>



- Ptaszek, G., Yuskiv, B., & Khomych, S. (2023). War on frames: Text mining of conflict in Russian and Ukrainian news agency coverage on Telegram during the Russian invasion of Ukraine in 2022. *Media, War & Conflict*, 17(1), 41–61. <https://doi.org/10.1177/17506352231166327>
- Robertson, G. (2022). Protest, platforms, and the state in the Belarus crisis. *Post-Soviet Affairs*, 38(1–2), 146–149. <https://doi.org/10.1080/1060586X.2022.2037196>
- Roth, C., & Cointet, J. P. (2010). Social and semantic coevolution in knowledge networks. *Social Networks*, 32(1), 16–29. <https://doi.org/10.1016/j.socnet.2009.04.005>
- Roth, C. (2013). Socio-semantic frameworks. *Advances in Complex Systems*, 16(4&5), 1350013. <https://doi.org/10.1142/S0219525913500136>
- Rudling, P. A. (2017). “Unhappy Is the Person Who Has No Motherland”: National Ideology and History Writing in Lukashenka’s Belarus. In: J. Fedor, M. Kangaspuro, J. Lassila, & T. Zhurzhenko (Eds.) *War and Memory in Russia, Ukraine and Belarus* (pp. 71–105). Palgrave Macmillan Memory Studies. Palgrave Macmillan. [https://doi.org/10.1007/978-3-319-66523-8\\_3](https://doi.org/10.1007/978-3-319-66523-8_3)
- Rust, M. (2022). Belarusian political elites: New, imagined, lost? *New Eastern Europe*, 6(54) 164–170.
- Silitski, V. (2006). Belarus: Learning from defeat. *Journal of Democracy*, 17(4), 138–152. <https://dx.doi.org/10.1353/jod.2006.0073>
- Slobozhan, I., Brik, T., & Sharma, R. (2023). Differentiable characteristics of Telegram mediums during protests in Belarus 2020. *Social Network Analysis and Mining*, 13(1), 19. <https://doi.org/10.1007/s13278-022-01002-7>
- Sulzhytski, I. (2022). Opposition as “A Mould on the Fatherland”: Hate Speech and Grassroots Telegram Propaganda in Belarus. *Journal of Belarusian Studies*, 12(1–2), 5–34. <https://doi.org/10.30965/20526512-12350018>
- Świerczek, M. (2022). (2022). The role of provocative-deceptive actions in neutralizing the Belarusian opposition. *Internal Security Review*, 26(14), 301–336. <https://doi.org/10.4467/20801335PBW.21.043.15703>
- Szostek, J. (2018). The mass media and Russia’s “Sphere of Interests”: Mechanisms of regional hegemony in Belarus and Ukraine. *Geopolitics*, 23(2), 307–329. <https://doi.org/10.1080/14650045.2017.1402298>
- Tao, W., & Peng, Y. (2023). Differentiation and unity: A Cross-platform Comparison Analysis of Online Posts’ Semantics of the Russian–Ukrainian War Based on Weibo and Twitter. *Communication and the Public*, 8(2), 105–124. <https://doi.org/10.1177/20570473231165563>
- Torichnyi, V., Biletska, T., Rybshchun, O., Kupriyenko, D., Ivashkov, Y., & Bratko, A. (2021). Information and propaganda component of the Russian Federation hybrid aggression: conclusions for developed democratic countries on the experience of Ukraine. *Trames: A Journal of the Humanities and Social Sciences*, 25(3), 355–368. <https://doi.org/10.3176/tr.2021.3.06>
- Wilson, A. (2021). *Belarus: the last European dictatorship*. Yale University Press. <https://doi.org/10.2307/j.ctv1h7zmxs>
- Wijermars, M., & Lokot, T. (2022). Is Telegram a “harbinger of freedom”? The performance, practices, and perception of platforms as political actors in authoritarian states. *Post-Soviet Affairs*, 38(1–2), 125–145. <https://doi.org/10.1080/1060586X.2022.2030645>
- Yurkova, O. (2018). Ukraine: At the Forefront of Russian Propaganda Aggression. *The SAIS Review of International Affairs*, 38(2), 111–124. <https://doi.org/10.1353/sais.2018.0021>

---

ROK SMRDELJ,\* ROMAN KUJAR\*\*  
& MONIKA KALIN GOLOB\*\*\*

Intersections. EEJSP  
11(1): 111–138.  
<https://doi.org/10.17356/ieejsp.v11i1.1220>  
<https://intersections.tk.hu>

## Expanding Boundaries: “Gender Theory” and the Twitter (X) Debate on Gender-Sensitive Language Use in Slovenia

---

\* [\[rok.smrdelj@ff.uni-lj.si\]](mailto:rok.smrdelj@ff.uni-lj.si) (University of Ljubljana, Faculty of Arts)  
\*\* [\[roman.kuhar@ff.uni-lj.si\]](mailto:roman.kuhar@ff.uni-lj.si) (University of Ljubljana, Faculty of Arts)  
\*\*\* [\[monika.kalin-golob@fdv.uni-lj.si\]](mailto:monika.kalin-golob@fdv.uni-lj.si) (University of Ljubljana, Faculty of Social Sciences)

### Abstract

This study examines the transition of the “gender theory” discourse from the realm of political and activist circles to the general population, focusing on the 2018 Twitter (X) debate in Slovenia regarding the gender-sensitive language policy at the Faculty of Arts, University of Ljubljana. Through a mixed-methods approach combining social network analysis and linguostylistic and critical frame analyses, this paper explores the dynamics and implications of this discourse shift. Our analysis identifies distinct user communities with primarily right-leaning political affiliations, revealing how these groups navigate and emphasize various themes related to gender-sensitive language, often broadening the debate to include new topics. The concept of “gender theory” emerges as a pivotal element, serving as a unifying thread that transforms discussions into ideological battlegrounds, thereby creating an “echo chamber” effect that sidelines opposing viewpoints. The findings underscore the colonizing effect of the “gender theory” discourse, which not only diverts attention to new ideological issues but also has a de-democratizing impact by constraining the range of acceptable debate and excluding divergent perspectives. This paper corroborates the initial premise that the discourse on “gender theory” has evolved, acquiring a broader societal and ideological dimension that challenges the inclusivity of public debate.

**Keywords:** gender theory, gender ideology, gender-sensitive language, anti-gender mobilizations, Twitter, social network analysis

## 1 Introduction

In the past few years, we have witnessed a proliferation of anti-gender campaigns organized by neoconservative groups that place the concept of gender at the roots of their anti-feminist and anti-LGBT campaigns. While initially focused primarily on opposing marriage equality, anti-gender mobilizations in Slovenia now also target sexual and reproductive rights, sex education in schools, transgender rights, gender liberalism, knowledge production and similar issues (Kuhar & Paternotte, 2017; Graff & Korolczuk, 2022). These issues are grouped under the umbrella of “gender theory” or “gender ideology,” serving as

---

rallying points for unifying diverse actors. The term “gender theory” functions as a “symbolic glue” (Kováts & Pöim, 2015) that enables “discourse coalition” (Edenborg, 2023) and “opportunistic synergy” (Graff & Korolczuk, 2022) among diverse actors. It provides them with a mobilization tool to represent a wide range of social issues. In this sense, as Paternotte (2023) puts it, anti-gender mobilizations have evolved beyond the Catholic Church’s control, their primary creator. Consequently, it is imperative to explore the diverse manifestations of gender ideology as a means of mobilization that transcends the initial boundaries of the anti-gender political agenda.

Slovenia serves as a compelling example of how “gender theory”<sup>1</sup> has gained widespread influence. Notably, Slovenia witnessed the culmination of this mobilization with legislative referenda in 2012 and 2015, resulting in the rejection of marriage equality. The early emergence of Slovenian mobilization in 2009 gave it a temporal advantage, allowing it to gain political experience, mobilize supporters, and achieve its first referendum victory ahead of similar mobilizations elsewhere. By utilizing enhanced tactics, building a robust support network, and establishing transnational connections, anti-gender actors solidified their position as significant political actors. While “gender theory” was not initially the central focus of their actions, the second referendum in 2015 provided an opportunity for anti-gender actors to legitimize it as a representation of their opposition to gender and reproductive rights, fundamentally shaping the public sphere (Kuhar, 2017, p. 224).

Both referenda were a major political victory for anti-gender actors in Slovenia, but the subsequent decision of the Constitutional Court in 2022, which legalized marriage equality and extended adoption rights to same-sex couples, undermined the efficacy of their efforts (Kuhar, 2023; Smrdelj & Pajnik, 2022). Nevertheless, it is important to note that the idea of “gender theory” has taken root, and its dissemination and use persist. The apocalyptic idea of a supposed plan for a cultural Marxist revolution to destroy masculinity, femininity, and family continues to take hold. Since 2015, “gender theory” has permeated the Slovenian public sphere and has become an unavoidable and accepted viewpoint in discussions on gender-related issues. Since 2015, there has scarcely been a public debate in Slovenia on sexual and reproductive rights in which “gender theory” as a discursive frame has not been one of the dominant viewpoints. This indicates a pervasive naturalization of “gender theory” in public debates on these matters.

In this article, we start from the premise that in recent years there has been a notable transition in the use of the “gender theory” framework from political and activist actors to the general population in Europe and beyond. “Gender theory” has evolved into a discursive framework for public deliberation, used not only among anti-gender activists, but also among the general public, media outlets, politicians, and other stakeholders (Paternotte, 2023).

<sup>1</sup> Slovenian anti-gender actors use the term “gender theory” to encompass what is referred to as “gender ideology” in other countries. In either instance, the term is used as an “empty signifier” (Mayer & Sauer, 2017). In this text, we use both terms synonymously, as we simultaneously refer to the broader transnational context and the specific Slovenian context in which anti-gender actors, and correspondingly the general public, use the term “gender theory”.

To test this claim in the context of Slovenia, we conduct quantitative and qualitative analyses of the 2018 Twitter (X)<sup>2</sup> debate surrounding the adoption of a gender-sensitive language decision by the Faculty of Arts, University of Ljubljana. The decision mandates the use of feminine grammatical gender to represent all genders in the internal regulations of the Faculty in the next three years, followed by the alternate use of masculine and feminine grammatical genders as neutral. Our study diverges from previous studies on anti-gender discourse by focusing on a case that was not the specific “target” of anti-gender actors in Slovenia. In contrast, this particular case, specifically chosen as it occurred three years after the introduction of “gender theory” into the Slovenian public sphere during the referendum campaign, aims to illustrate the transformation of the structure of public debates in Slovenia brought about by “gender theory.” Notably, the decision of the Faculty of Arts sparked a scandal in Slovenia, fuelled partly by a misinterpretation that quickly propagated through mass and social media. This misinterpretation falsely claimed that male individuals would be addressed in the feminine form at the Faculty of Arts (Kuhar & Antić Gaber, 2022). Major mass media extensively reported on the decision, triggering a lively, albeit predominantly offensive, Twitter debate, which, we believe, was framed by the “argument” of “gender theory.”

The Faculty’s Senate’s decision to incorporate the feminine grammatical gender as a neutral option in the internal regulations of the Faculty of Arts at the University of Ljubljana sparked a range of scholarly debates (Gorjanc & Stabej, 2019; Mikić & Kalin Golob, 2019; Šorli, 2019; Kuhar & Antić Gaber, 2022; Smolej, 2022), roundtable discussions (Gorjanc et al., 2019), and conferences. However, none of these scholarly activities have explored the debate surrounding gender-sensitive language use on Twitter through the lens of “gender theory.”

While Twitter may not be the most popular social media platform in Slovenia,<sup>3</sup> its selection for our analysis was based on its strong focus on current political debates, a feature less pronounced on Facebook, which is generally used for maintaining social connections with individuals known offline. Additionally, Twitter’s “architecture” allows for the observation of communication activities without the necessity of logging in or being directly connected to other users (Jesenšek et al., 2021; Verweij, 2012). This aspect, along with Twitter’s emphasis on socio-political discussions, significantly influences public debate and mass media reporting (Moon & Hadley, 2014), making it an influential “agenda-setting” medium, despite having fewer users than Facebook or Instagram. Thus, our choice of the medium for this study was guided not by user numbers but by the platform’s open “media logic” and its broader impact on the public sphere. Moreover, Instagram, with its focus on image-centric communication, was not considered suitable for our study, which concentrates on the linguistic aspects of the “gender theory” discourse. Last but not least,

---

<sup>2</sup> As this social medium was still known as Twitter at the time of the study, we will retain this naming throughout the article.

<sup>3</sup> According to the MEDIA+ survey by Valicon (2020), there were 198,000 Twitter users aged 16–74 in Slovenia in 2018, ranking it as the third most used social media platform after Facebook (940,500 users) and Instagram (397,000 users).

as we focus on the “gender theory” discourse, it is important to consider that Twitter is also favoured by anti-gender actors, who are known to effectively utilize digital tools to further their agendas (Kuhar & Pajnik, 2020; Popič & Gorjanc, 2022).

With the aim of demonstrating how “gender theory” has flooded the 2018 Twitter debate on the gender-sensitive language decision implemented by the Faculty of Arts in Ljubljana, our article begins with an introduction to the theoretical framework and proceeds to outline the research questions, sample, and methodology. This is followed by the presentation and discussion of the results of the retweet network analysis, as well as linguostylistic and critical frame analyses of the most retweeted tweets.

## 2 “Gender Theory/Ideology” and Twitter: Studies and Concepts

Studies on Twitter and “gender theory/ideology” demonstrate a key role of social media in naturalizing and reproducing “gender theory” as a discursive frame in a wide array of critical political and social matters. In a study on Twitter in Spain, Puente et al. (2021) highlighted the role of “gender ideology”, and more specifically the term “feminazi”, in the articulation of a specific narrative of masculinity. Popič and Gorjanc (2022) examined Slovenian texts and media and pointed out Twitter’s significance in maintaining anti-gender ideology. Ozuzden and Korkut (2020) found Twitter’s facilitation of conservative, homophobic reactions in Turkey, rooted in “gender theory”. Similarly, Wallaschek et al. (2022) examined gender equality debates on Twitter during International Women’s Day 2021 in Germany, Italy, and Poland, revealing how actors opposing gender equality utilized gender ideology to undermine feminist causes. Martinsson and Ericson (2023) studied grassroots anti-gender mobilizations in Sweden during the pandemic and confirmed the relevance of Twitter in establishing anti-gender networks. Together, these studies provide valuable insights into the role of Twitter in shaping and perpetuating discourses surrounding “gender theory” and its impact on social and political contexts.

Our study contributes to this body of literature with a Slovenian case study on gender-sensitive language debates on Twitter. However, to substantiate the theoretical underpinnings of the “colonizing” nature of “gender theory,” it is essential to first consider a few central concepts regarding “gender theory” which account for its rapid dissemination across diverse domains and its transformative impact on the public sphere concerning gender and related issues.

Firstly, the concept of the “symbolic glue” (Kováts & Pöim, 2015) asserts that “gender theory” serves as a unifying force, bringing together actors who may not have previously collaborated or shared identical ideological perspectives on specific topics, yet find common ground through the “gender theory” discourse and rhetoric. Graff and Korolczuk (2022) call such collaborations and alliances “opportunistic synergy”. The adoption of the anti-gender discourse is used as a means to bolster the popularity of these actors, rendering them the true defenders of the “common people” against the “corrupted elites”.

Secondly, the concept of the “empty signifier” (Mayer & Sauer, 2017) explains the effectiveness of “gender theory” by its ability to be filled with varied content due to its lack of specificity. The mobilizing effect is achieved by the “flexibility” of the meaning

of “gender theory” or “gender ideology”, which can easily change according to the needs of a specific mobilization and the audience it addresses. In this sense, “gender theory” is highly adaptable and not constrained by any specific subject matter.

Lastly, the “ideological code” (Smith, 1999; Paternotte, 2023) sheds light on how frequent and spontaneous repetition of a particular signifier in the public sphere conceals its intellectual origins, elevating it to an unquestionable truth. Consequently, individuals adopt and propagate this signifier without necessarily being aware of its origin, further bolstering its legitimacy and perceived accuracy as a reflection of reality.

To demonstrate the “colonizing” nature of “gender theory” on Twitter, we also employ some of the network theory concepts that inform our analysis. Firstly, we employ the concept of “central users,” which is based on the premise that not every user in an online debate possesses equal influence or visibility. While there are several ways to measure users’ influence on Twitter, such as the number of followers, likes, or retweets, our study specifically focuses on retweets. We consider retweets as a key indicator of influence, as they represent the “measure” of prominence and visibility of certain posts within the Twitter debate (Jesenšek et al., 2021). Put differently, the frequency of retweets serves as a measure of the support for and interest in a user’s content (Metaxas et al., 2021). Specifically, more retweets increase the visibility and distribution of a particular post in the online debate, resulting in its greater influence.

Employing “retweet” as a basic unit of measuring influence, we examine “centrality” by differentiating between the “in-degree” and “out-degree” measures of centrality. They are both crucial in studying directed graphs where edges have a clear direction from one node to another. Namely, the “in-degree” centrality measures the number of edges (retweets) originating from potentially different nodes (users) directed towards a node (user), indicating how frequently a user’s posts are retweeted. Conversely, the “out-degree” centrality refers to the number of edges (retweets) originating from a specific node (user) and directed towards other nodes (users), showing how often a user retweets others. However, in this study, our focus is only on the “in-degree” centrality measure to identify “central users”, as this measure indicates which users receive the most retweets and have the highest influence on the direction of the debate (Hansen et al., 2020a; Jesenšek et al., 2021). While identifying users’ activity (the “out-degree” centrality measure) is also a valuable aspect of social network analysis, our study does not explore this aspect as we are interested in the impact and reach of posts, rather than users’ activity.

Studying communication patterns in the digital media sphere is a crucial component of network theory (e.g. Arguedas et al., 2022; Kubin & von Sikorski, 2021; Van Bavel et al., 2021; Yarchi et al., 2021). For our study, the concept of “community” (Hansen et al., 2020b) is particularly important as it underlines how users engage with certain topics online. That is to say, interactions among users vary in frequency and content – some users communicate more frequently, while others do so less or not at all, and their opinions on issues may align or diverge. This variation allows for the algorithmic identification of communities, grouping users who interact more frequently into the same community (see the methodological part of our study). These are represented as clusters in network visualizations, where closer nodes (users) indicate more frequent communication. If network analysis uncovers distinct and fragmented communities, it enables further investigation into commu-



nication patterns and trends, for instance, the “echo chamber” effect (e.g. Arguedas et al., 2022) and “network polarization” (e.g. Kubin & von Sikorski, 2021; Van Bavel et al., 2021).

“Echo chambers” and “network polarization” are identified both visually and through content analysis of the users’ posts. The “echo chambers” effect occurs when users within the same community consistently reinforce each other’s viewpoints (e.g. Arguedas et al., 2022), a phenomenon often observed in the Twitter “retweets” that amplify messages within the same community (e.g. Smrdelj, 2024). Conversely, network polarization is observed when users across different communities engage in discussions from strongly opposing ideological viewpoints on a specific subject. It is evident through the segregation of communities in network visualizations and through the content analysis that reveals ideologically contrasting positions between these communities. It is important to point out that both communication patterns are understood in our study as a result of social media architecture, i.e. as a structural feature inherent in the logic of contemporary social media (e.g. Kubin & von Sikorski, 2021; Van Bavel et al., 2021).

By integrating insights from anti-gender studies and network theory, we examine and explore the “colonizing” nature, remarkable adaptability, and flexibility of “gender theory/ideology”. As a mobilizing tool, it has absorbed various topics, frameworks, and narratives, which have contributed to its transnational nature; it is adaptable to different national, cultural, religious, and political contexts. Lacking internal coherence, it can be moulded by different actors to serve diverse political projects (Paternotte, 2023). Based on this premise, our main theoretical argument posits that “gender theory” as a discursive framework has significantly reshaped public debate surrounding gender and related topics, leveraging the support of mass media and effectively utilizing social media platforms. This has been achieved by appropriating the discourse of progressive social movements, strategically invoking concepts such as human rights and democracy. This appropriation poses a challenge, as it creates a facade that makes it more difficult to contest the underlying ideologies and actions associated with gender “theory” (Smrdelj et al., 2021). Consequently, the rhetoric and discourse surrounding “gender theory” play an active role in the contemporary processes of de-democratization (Lombardo et al., 2021). Their swift and effective dissemination in the public sphere reinforces the dominance of anti-gender voices while simultaneously contributing to polarization within the public sphere, in which dialogue is limited to *pro et contra* positions. This impedes reasoned debate on a range of important political and social issues, as an inclusive and nuanced debate is made impossible by the vociferous anti-gender usurpation of public discourse.

### 3 Data Collection Methods and Sample Selection

The quantitative part of our research focuses on the identification of distinct Twitter communities and their size in terms of users and retweets. We also identify the central users in each community. The qualitative part of the study focuses on the specific linguostylistic elements and frames that characterize the most retweeted tweets within aforementioned communities.

To carry out both parts of the study, we obtained data on January 19, 2023, utilizing the Twitter API Academic Access.<sup>4</sup> Our sample consists of all Slovenian Twitter posts between May 27, 2018, and October 30, 2018. The search terms used to identify Twitter posts within the specified timeframe were chosen through the following procedure. Firstly, we derived most of our search terms from the vocabulary related to the resolution adopted by the Faculty of Arts: “spol” (gender/sex), “moški spol” (male gender), “ženski slovnični spol” (feminine grammatical gender), “katerikoli spol” (any gender), “senat FF” (Senate of the FF<sup>5</sup>), “sklep senata FF” (Decision of the Senate of the FF), “sklep filofaksa” (Decision of the Filofaks<sup>6</sup>), “filofaks” (Filofaks), “vodstvo filozofske fakultete” (Leadership of the Faculty of Arts), “vodstvo FF” (Leadership of the FF), “Roman Kuhar” (Roman Kuhar<sup>7</sup>), and “dekan Kuhar” (Dean Kuhar). These terms are considered neutral, as they were used by different public actors irrespective of their stance on the proposal or political affiliation, as shown by existing research (Kuhar & Antić Gaber, 2021). Secondly, we included scholarly terminology from existing Slovenian research that specifically addressed the “linguistic” dimension associated with the Faculty of Arts measure (Šorli, 2019). In this instance, we selected the following search terms: “občutljiva raba jezika” (sensitive use of language), “generična raba” (generic use), and “nevtralen” (neutral). Similarly, as in the previous case, these three terms also have a neutral meaning. Thirdly, to ensure coverage of the “gender theory” discourse in the Twitter debate on gender sensitive language use, we also included less neutral and more “anti-gender” terms such as “teorija spola” (gender theory), “radikalni feminizem” (radical feminism), and “LGBT ideologija” (LGBT ideology). In the selection of these words, we aimed to encompass not only “neutral” and “scholarly” search terms but also the terms linked to the ideological discursive foundation of those opposing gender-sensitive language use. Finally, we searched for posts in Slovenian using listed keywords across all six grammatical cases. This approach aimed to capture the Slovenian debate, given that Slovenian users primarily communicate in their native language (Evkoski et al., 2021).

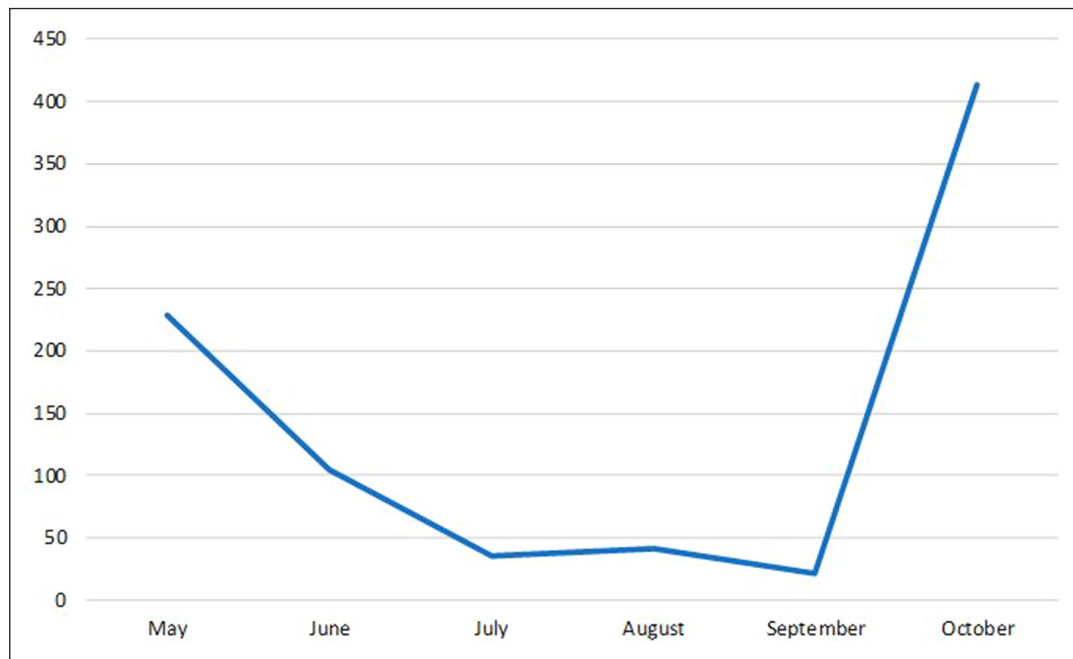
The onset of the heated public debate about the Senate’s decision can be traced back to May 26, 2018, when the first newspaper article regarding the decision was published in the mass media (Ahačič, 2018). The debate continued fervently until the end of June, experienced a lull during the summer, and then reemerged on October 23, 2018, when a roundtable discussion on gender-sensitive language in the internal regulations of the Faculty of Arts at the University of Ljubljana was organized. To ensure a comprehensive coverage of the debate, we sampled Twitter posts from May 27, 2018 (the day following the first newspaper article on the decision) until October 30, 2018, including the week following the roundtable. Graph 1 illustrates the fluctuating intensity of the debate, depicting the monthly numbers of tweets throughout the analysed period.

<sup>4</sup> In obtaining the Twitter data, we collaborated with Uroš Godnov from the Institute for Data Quality (Slovenian: Inštitut za kakovost podatkov, d.o.o.).

<sup>5</sup> FF is the official abbreviation for the Faculty of Arts, University of Ljubljana.

<sup>6</sup> Popular unofficial name of the Faculty of Arts, University of Ljubljana.

<sup>7</sup> The name of the dean of the Faculty of Arts at the time of the adoption of the decision.



Graph 1 Number of Twitter Posts per Month

Source: Generated by the authors

The initial sample included 1,403 posts. After conducting data cleanup, all the posts that did not align with our research topic were removed, resulting in the deletion of 559 posts. Consequently, the refined sample comprised 844 posts generated by 364 users (Table 1). However, while our dataset may be smaller in comparison to some network analyses in Slovenia, such as the one by Jesenšek et al. (2021) in the study on the European Parliament elections or by Smrdelj (2024) in the study on the migrant “crisis”, it is important to note that our sample effectively represents the “population”, encompassing all Slovenian users and posts related to this topic in the specified time period. The reason for the comparatively smaller sample size compared to Twitter debates in analogue studies is that existing Twitter studies often focus on the topics that occupied a dominant position in the media agenda at a given time. On the other hand, our study intentionally focuses on the public issue that did not have a dominant media coverage. Since the issue of gender-sensitive language use was not a dominant topic on the media agenda at a given time (like elections, migrant “crisis” or marriage equality referendum if we focus on the topics that are related to the anti-gender agenda), it is critical for illustrating the “colonizing” nature of “gender theory” in the public sphere. Furthermore, as we explain in the following section, our study is mostly qualitative, focusing on the most retweeted tweets, while the quantitative dimension serves primarily as the foundation for selecting the most retweeted tweets.

**Table 1** Analysis Sample

<b>Data</b>	<b>No.</b>
Posts	844
Tweets	399
Retweets	445
Faculty Senate Decision Retweets	177
Other retweets	268
Users	364
Users who retweeted and were retweeted by	255

*Source:* Generated by the authors

#### **4 Methodological Approach: Utilising Social Network Analysis, Linguostylistic Analysis, and CFA**

To construct and analyse the retweet network, we employed social network analysis methods, allowing us to ascertain the central users within the network, identify the communication relations among them, and examine the content of these relations (Ferra & Nguyen, 2017; Hansen et al., 2020b). Following Al-Rawi (2019), Jesenšek et al. (2021) and Smrdelj (2024), we created a retweet network, which is expounded upon in the subsequent section alongside our analysis findings.

In addition to the social network analysis, we employed a linguostylistic analysis (Jeffries & McIntyre, 2018) and a CFA (Verloo, 2005). By merging quantitative (social network analysis) and qualitative (linguostylistic analysis, CFA) methodologies, we gain a multifaceted perspective on network characteristics, facilitating a deeper understanding of our data.

Linguostylistic analysis involves the examination of the linguistic aspects pertaining to the construction of meanings within a given topic and it enables to “approach the explicitly persuasive aspects of style as linguistic phenomena” (Jeffries & McIntyre, 2018, p. 8). This approach focuses on linguistic choices and attitudes in the formation of beliefs and presupposes that every lexical item evokes a frame. “Linguistically induced evocation of frames in political thought not only guides our comprehension of issues at given moment but, furthermore, strengthens those frames in our minds” (Wehling, 2020, pp. 138–139). Linguostylistics studies phenomena concerning the systemic aspect and use of language, and analyses the use of linguistic means in texts and their systemic characteristics (Čechová et al., 2008, pp. 44–51). When analysing the language of (social) media texts, the aims of the examination are diverse, ranging from identifying linguistic changes, development, general use and the emergence of new words to studying how topical themes are articulated by a given medium to inform and influence the public. The first aspect points

to the influence of media language on the use and codification of the standard language, and the second one accounts for the selection of linguistic means as a tool of influence and valuation (Kalin Golob & Grizold, 2017). This study addresses this second aspect and seeks to determine how the users reacted to a delicate issue to influence the general public.

On the other hand, the CFA (Verloo, 2005) adopts a discursive approach to explore the underlying norms, beliefs, and perceptions conveyed in the texts. This method operates on the assumption that each text consists of one or more frames that are characterized by specific markers (Dombos et al., 2012). These markers serve as responses to “sensitization questions” (Verloo & Lombardo, 2007, p. 35), which assist in identifying various aspects of the text, such as the definition of the problem (i.e. diagnosis), proposed solutions (i.e. prognosis), the actors responsible for the problem (i.e. active actor), and those affected by it (i.e. passive actor).

We treated each tweet as a problem statement, potentially encompassing both a problem definition and an implicit solution. Such a statement may also involve identifying the actor(s) causing the problem and the actor(s) victimized by it. These components serve as markers upon which frames are constructed. Alongside the four dimensions mentioned earlier (problem definition, solution given, active actor, passive actor), we also examined the underlying norms expressed in each tweet. This refers to the position from which a particular statement is made, reflecting normative viewpoints. Based on these five dimensions, we constructed frames, which Verloo defines as “organizing principles that transform fragmented or incidental information into a structured and meaningful problem, implicitly or explicitly incorporating a solution” (Verloo, 2005, p. 20). Therefore, frames represent the explanatory or cognitive structures utilized by individuals to interpret, shape, and possibly reinforce a specific social reality.

Combining CFA and linguostylistic analysis enabled us not only to uncover the underlying socio-political perspectives and ideological stances embedded in the discourse, but also to understand how language is used to evoke emotions and engage the audience. Employing both methodological approaches, we managed to shed light on the broader social and cultural context as well as on the specific rhetorical devices, linguistic patterns, and communicative strategies employed by influential users to capture attention and shape public opinion.

The subsequent analysis is structured into two sections. The first section entails the quantitative findings derived from the retweet network construction. The second section delves into the most retweeted tweets, providing insights into their linguostylistic features and the underlying discursive frames.

## 5 Analysis Results

### 5.1 Retweet Network Creation and Visualization

To create a retweet network, we built a directed, non-weighted graph between the Twitter users who retweeted each other. The direction indicates information flow. The retweet network was built utilising the following stages. Firstly, in R software (version 4.2.2.), we selected all the users involved in generating retweets, including those who authored the retweeted posts and those who retweeted them. This process yielded 255 users, who generated 455 retweets, which signified the relations between users. Subsequently, we imported the data into Gephi (version 0.9.2.),<sup>8</sup> which provided 394 “unique edges” (Hansen et al., 2020b) between users. A “unique edge” signifies a relationship between two users, regardless of the number of the retweets exchanged between them. For instance, if user A retweeted user B’s posts five times, we have five retweets (and hence, five relations) between them, but only one “unique edge”. Thus, the initial count of “455” represents the total number of relations corresponding to the overall number of retweets, whereas the count of “394” represents the “unique edges” that served as the foundation for generating the network visualization in Gephi.

Following the identification of the “unique edges” between users, distinct communities within the retweet network were detected using the Louvain algorithm. This algorithm relies on the identification of shared communication patterns among network users to delineate communities (Blondel et al., 2008). The algorithm identified 24 distinct communities within the network. To streamline our analysis, we focused on the seven largest communities, collectively encompassing 76 per cent of all users within the retweet network. The primary reason for excluding some communities from the analysis was their relatively small user base and lower activity level. Unlike the seven largest communities, these smaller groups had fewer posts and significantly less engagement, as evidenced by a lower number of retweets. This limited engagement meant they did not significantly contribute to the broader debate, leading to their omission from our consideration. Detailed characteristics of each community in the retweet network are presented in Table 2.

We selected the number of users as a starting point for determining the size of each community (Table 2, second column). Additionally, the third column of the table specifies the count of the retweets generated by users in a particular community, whereas the fourth column provides an overview of the number of the “unique edges” connecting them. Moreover, the final two columns in Table 2 present summaries of the key framings (CFA) and key linguostylistic strategies. Since the same linguostylistic strategies appear more or less frequently in all communities, they are not categorized by individual community in the last column but are applicable to all communities.

The retweet network of the seven largest communities was visualized using Gephi (Figure 1). The network visualization consists of 195 nodes (users) and 346 “unique edges”.

<sup>8</sup> During our work with Gephi, we collaborated with Luka Jesenšek, whose expertise greatly contributed to visualizing the retweet network and identifying communities using the Louvain algorithm.

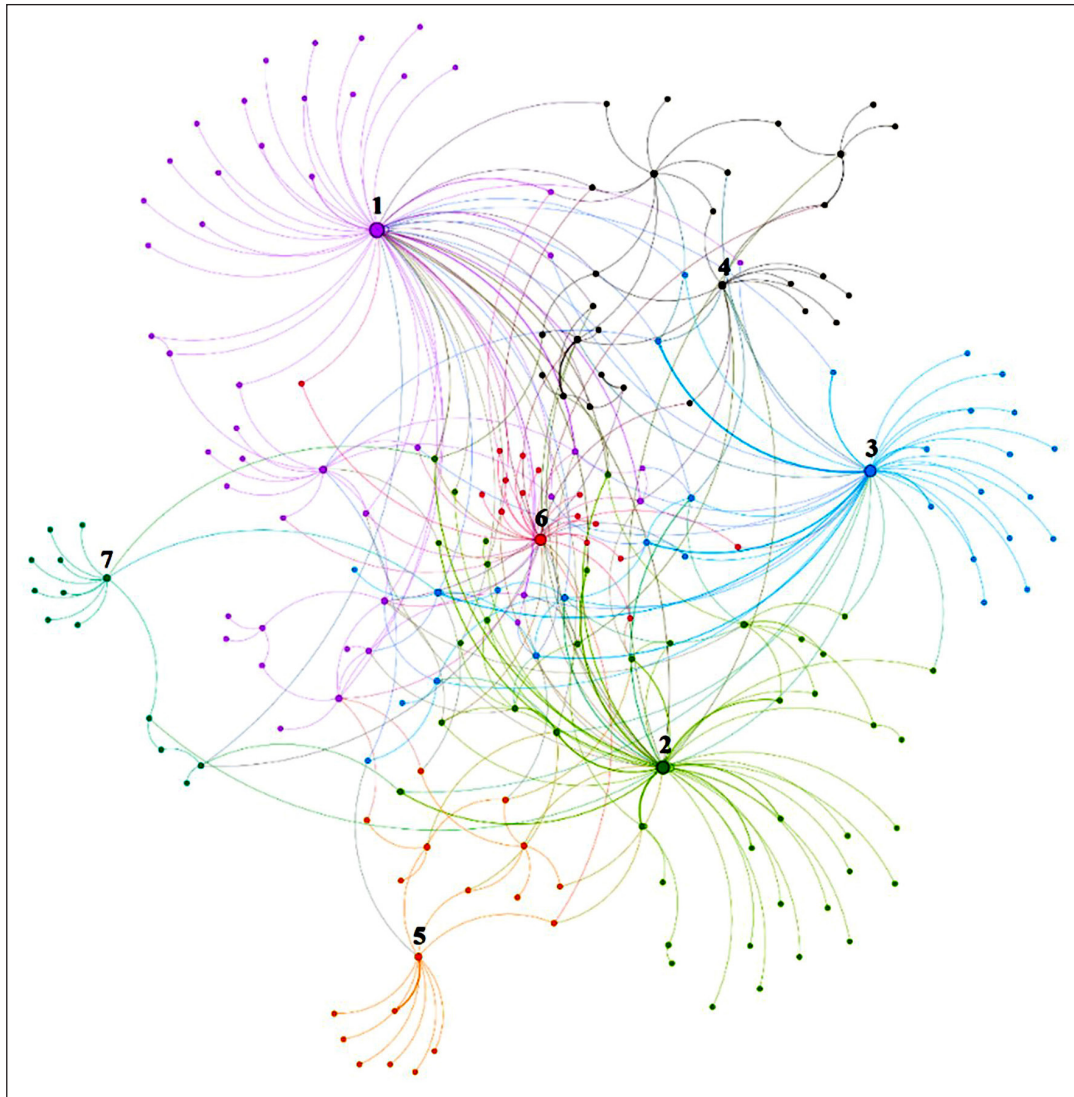


Table 2 Retweet Network Communities

Communities	Users	Retweets	Unique edges	Key Framings (CFA)	Key Linguostylistic Strategies
Community 1 (Education Guardians)	47	98	92	Educational indoctrination threat frame	Standard language; sub-standard as a tool of valuation  Non-specialized use of terms with a pejorative connotation  Use of metaphors and irony  Use of neologisms  The exclamation mark for emotionally charged sentences to express indignation  Use of slogans in English language  Use of puns  Use of defamations  Use of hate speech
Community 2 (Tradition Defenders)	43	108	76	Traditional gender norms restoration frame	
Community 3 (Dean’s Critics)	30	71	61	Dean-centric “gender theory” frame	
Community 4 (Cultural Marxism Challengers)	28	51	45	Cultural Marxism frame	
Community 5 (Linguistic Purists)	18	29	21	Linguistic intrusion of the “gender theory” frame	
Community 6 (Satire Supporters)	16	21	36	Satirical critique of the “gender theory” frame	
Community 7 (Satire Supporters)	13	14	15		
Community 8–24	60	53	48	/	/
Sum	255	445	394	/	/

Source: Generated by the authors

In Figure 1, the distinct communities within the retweet network are represented by numbers and corresponding colours: Community 1 (purple), Community 2 (green), Community 3 (blue), Community 4 (black), Community 5 (orange), Community 6 (pink), and Community 7 (green-blue). Notably, the users who engaged in more frequent communication are positioned in closer proximity to one another, whereas users with less frequent communication are situated at greater distances from each other within the visualization. Our quantitative analysis has identified distinct communities within the retweet network. However, retweet network visualization alone cannot sufficiently reveal specific communication patterns, such as the "echo chamber" effect or network "polarization" among users. To ascertain this, we integrate the findings from network visualization with outcomes from the CFA and linguostylistic analysis, detailed in the subsequent section.



**Figure 1** Visualization of the seven largest retweet network communities

*Source:* Generated by the authors

Expanding on the “central users” concept introduced in the second section, we determined central users by analysing the number of retweets they received from other users (“in-degree” centrality) and simultaneously their “unique edges”, as detailed in Table 3. To provide an in-depth understanding of the network’s characteristics, we introduced two additional attributes for each user: “category” and “political affiliation”. Under “category,” we assigned a specific status to each user (e.g. media, politician). For users whose identity was unknown, we categorized them as either “publicly unknown user” or “anonymous user.”

This classification was based on their profile information: users with a first and last name and/or a personal profile photo, yet not publicly recognizable, were labelled as “publicly unknown.” Conversely, those with a pseudonym and no personal photo were classified as “anonymous users.” Moreover, regarding the “political affiliation”, it should be noted that all the posts by central users in the sample align with right-wing political ideology. Therefore, it can be asserted that a network with a right-leaning political orientation dominates the discussion. However, in Table 3’s last column, we only identified the political affiliations of publicly known users with known right-wing affiliations. We have not determined the affiliations of publicly unknown or anonymous users due to the inability to ascertain their personal ideological and political orientations.

Table 3 Central Users

Communities	Central Users	Retweets	Unique Edges	Category	Political Affiliation
Community 1 (Education Guardians)	@Nova24TV	72	59	Media	Right-wing
	@VsenoviceInfo	15	15	Media	Right-wing
	@Max970	7	7	Publicly Unknown User	/
	@PetraCernetic	4	4	Publicly Unknown User	/
	@Miran7777	3	3	Anonymous User	/
	@BenjaminNatanja	1	1	Anonymous User	/
	@MareAndi	1	1	Publicly Unknown User	/
	@MStrovs	1	1	Politician	Right-wing
	@Margu501	1	1	Publicly Unknown User	/
	Sum	105	92	/	/
Community 2 (Tradition Defenders)	@Anti_MigrantSLO	57	42	Media	Right-wing
	@Missspela	10	10	Publicly Unknown User	/
	@SnsStranka	11	6	Political Party	Right-wing
	@ifi_genija	7	6	Publicly Unknown User	/
	@SSN_pr	6	5	Political Party	Right-wing
	@BozoPredalic	2	2	Politician	Right-wing
	@MervicVanda	1	1	Publicly Unknown User	/
	@AnaKlaraT	1	1	Publicly Unknown User	/
	@Diphenbachia	1	1	Anonymous User	/
	@MarijaSoba	1	1	Publicly Unknown User	/
	@KerMiNiVseeno	1	1	Anonymous User	/
	Sum	98	76	/	/

Table 3 (continued)

Communities	Central Users	Retweets	Unique Edges	Category	Political Affiliation
Community 3 (Dean's Critics)	@RomanVodeb	50	40	Self-proclaimed psycho-analyst	Right-wing
	@DarinkaVrabi	6	6	Publicly Unknown User	/
	@petra_jansa	5	5	Journalist	Right-wing
	@JureUmnik	5	5	Publicly Unknown User	/
	@mojcav1	3	3	Publicly Unknown User	/
	@twiitiztok	2	2	Publicly Unknown User	/
	Sum	71	61	/	/
Community 4 (Cultural Marxism Challengers)	@BernardBrscic	16	16	Politician	Right-wing
	@Muhabc1	11	11	Publicly Unknown User	/
	@RadioOgnjisce	8	6	Media	Right-wing
	@Tomltoml	6	5	Anonymous user	/
	@ATBeatris	4	3	Publicly Unknown User	/
	@krtmen	3	3	Anonymous user	/
	@ZigaTurk	1	1	Politician	Right-wing
	Sum	49	45	/	/
Community 5 (Linguistic Purists)	@PlusPortal	13	11	Media	Right-wing
	@Alex_Zamuda	6	6	Politician	Right-wing
	@cnfrmstA	4	4	Anonymous user	/
	Sum	23	21	/	/
Community 6 (Satire Supporters)	@Casnik	36	36	Media	Right-wing
	Sum	36	36	/	/
Community 7 (Satire Supporters)	@Peterstrov	11	11	Publicly Unknown User	/
	@SamoGlavan	2	2	Musician	Right-wing
	@Ajitamxy	1	1	Publicly Unknown User	/
	@Demokracija1	1	1	Media	Right-wing
	Sum	15	15	/	/
	Total Sum	397	346		

Source: Generated by the authors

All seven communities were dominated by right-wing central users. Community 1 represents *Nova24TV*'s network, which is a leading right-wing TV station in Slovenia. The primary central user in Community 2 is an anonymous user whose username indicates their opposition to immigration. Moreover, the profile of the extra-parliamentary, right-wing, Eurosceptic, populist SNP (Slovenian National Party) can be identified among the prominent users in this community. Community 3 revolves around the controversial psychoanalyst Roman Vodeb, who promotes "gender theory" and other exclusionary views under the guise of psychoanalytic science. Community 4 centres on the far-right politician Bernard Brščič, the leader of the extra-parliamentary Homeland League, who is known for his inflammatory attitudes regarding social minorities. Community 5 consists of the conservative online news portal *Portal Plus*. Similarly, Community 6 revolves around the conservative online news portal *Časnik*. Finally, Community 7 was formed around Peter Štrovs. Among the central users in Community 7 is the right-leaning weekly *Demokracija*, which publishes Štrovs' blog posts.

## 5.2 Linguostylistic Analysis and CFA of the Most Retweeted Tweets

Through network visualization, we can gain insight into the prevailing communication patterns among users. However, to discern the distinctive characteristics and thematic aspects of the discussions within each community, we also conducted a detailed linguostylistic analysis and CFA of the most retweeted tweets within them. Since all seven communities use similar linguistic strategies, the linguostylistic analysis was conducted at the level of all seven communities (the features are listed in Table 2), while the CFA was conducted separately for each community.

### 5.2.1 Linguostylistic analysis

A linguostylistic analysis has revealed that most of the tweets were written in a standard language, with non-standard language appearing as a stylistically marked category, playing the role of evaluating persons or actions, often with a derogatory connotation. In some examples, it also takes on an ironic tone. However, the replies to the most retweeted tweets exhibit less adherence to standard language, incorporating dialectal variations, vulgarity, and even expressions of hostility.

A number of notable categories of lexemes appear in the most retweeted tweets. Among them, appellations that function as pseudo-terms stand out, such as "gender theory", "radical feminism", "feminist ideology", "cultural Marxism", and "ideological activism". The non-specialized nature of these terms is evident in the fact that they have a pejorative connotation, which becomes apparent in the context. In all seven communities, the core phrases that describe the Faculty's action are theory/gender ideology, LGBT agenda/propaganda, LGBT-feminist ideology, and leftist/(cultural) Marxist ideology.

Another category of lexemes consists of adjectives like "sick" and "perverted," as well as adjectives that are predominantly emotionally tainted with a derogatory connotation.

Among them are adjectives such as “gay dean”, “misguided ‘gender theory’”, “forced use of new pronouns”, “effeminate generation”, “stupid idea”, “confused fairy”, and “leftist-faggot delusions”.

The tweets also include stylistically marked colloquial and expressive verbs and nouns with derogatory and potentially offensive connotations, as well as expressive designations, such as “delusions”, “charlatans” “invaders”, “faggotism”, “foolish women”, “rags”, “primitives in power”, “globalist”, “freemason”, and “gravedigger of manliness”.

Along with the aforementioned noticeable lexemes, punctuation – specifically the exclamation mark – is a linguistic choice that significantly frames the tweets in our analysis. The exclamation mark, according to the Slovenian spelling convention, ends emotionally charged sentences. In analysed tweets it is used to express indignation (“*They want to create a generation of effeminate men through ‘gender theory’!*”; “*An attack by a perverted and degenerate leftist ideology on the Slovenian language and normality!*”), underlines individual claims (“*LGBT-feminist ideology has no place in educational institutions!*”; “*No one will ever again enter the curricula of faculties with these leftist-feminist delusions!*”) or stands at the end of explicit appeals (“*Defend the family!*”; “*He should be fired, and not only from the dean’s post, from the Faculty!*”).

In an emotionally and stylistically marked usage in analysed tweets, the question mark is often used to end rhetorical questions (“*Are we just going to let them raise our children like this?*”) or to emphasize wonder (“*Is she the only one at the Department of Slovene studies?*”).

Ultimately, the linguostylistic analysis revealed that the phrases utilized in the tweets exhibit a negative connotation or prominently feature crass and vulgar expressions (e.g. “devil’s noose”, “kicking in the dark”, “pussy smoke”). Furthermore, the employment of metaphors within these tweets often deteriorates into unsuccessful or incomprehensible catachresis. Ironically, the writers defend the authentic, “real Slovenian” while using a slogan in the English language taken from an online collection of photographs (“Defend the nuclear family”). The selection of marked expressions – words with expressive value, ambiguous linguistic meanings, punctuation, and metaphors along with the style of the entire posts – plays a crucial role in assessing their influence function, namely, the persuasiveness of the linguistic means (Fairclough, 1994, p. 119).

## 5.2.2 CFA

In the quantitative phase of our analysis, we identified seven distinct user communities. A critical examination of the discourse frames utilized within these communities revealed the presence of two interconnected levels that characterize each community’s discourse. The first level can be referred to as the “root” or “foundation,” which remains consistent across all the identified groups. The second level represents the “stem” or substantive manifestation specific to each group’s discursive framework.

In our analysis, all the tweets share a common underlying theme of “gender theory” as the foundational concept upon which each tweet is constructed. It is important to note that “gender theory” serves as the contextual backdrop for our analysis, as the keyword itself guides our investigation. However, we aim to emphasize the contextual manifestation or discursive frames stemming from this shared foundation.



The most retweeted tweets in Community 1 are characterized and framed by the perceived threat of an ideology infiltrating the education system, targeting and allegedly indoctrinating innocent children. This discursive frame highlights the concerns surrounding the introduction of such an ideology into schools.

*Hungary finally bans the misguided gender theory at postgraduate level – brought there by the globalist Soros! (@Nova24TV)*

*Gay dean of the Faculty of Arts bans male gender, psychoanalyst Vodeb critical of him: “LG-BT-feminist ideology has no place in educational institutions!” (@Nova24TV)*

Frames in Community 2 emerge from a patriarchal standpoint that calls for the restoration of traditional gender norms and relations, presenting it as the normative and desirable state of affairs.

*Well done Hungarians! One of the few healthy nations in Europe. We in the SNS are also strongly against gender theory. Men and women should have equal rights, but should not be equal! (@Sns-Stranka)*

*Let us defend our women! This is what feminism should fight for. Not for a ban on showing pretty girls during football matches. Radical feminism is a disease. (@Anti\_MigrantSLO)*

The most retweeted tweets in Community 3 focus specifically on the situation at the Faculty of Arts, where the discursive frame revolves around arguments *ad personam*, positioning the Dean of the Faculty as the central agent of “gender theory”. Within this frame, “gender theory” is portrayed through an emphasis on the Dean’s personality characteristics, including alleged psychological traumas from childhood and his sexual orientation.

*Now the young men, driven by their testosterone, who have “sperm splashing out of their ears” should banish Dean Roman Kuhar – the gravedigger of manliness at the Faculty of Arts. (@RomanVodeb)*

*Kuhar is a latent militant gay who has ideologically and in line with the regime incorporated himself into the education system. Every gay man is created in a difficult primary environment/family, perhaps with a background of sexual abuse. A deficit in moral reasoning (Superego) is inevitable. (@RomanVodeb)*

Community 4 is characterized by political allusions that revolve around the ideological polarization between the left and the right. Specifically, it portrays “gender theory” as an offspring of left-wing woke individuals. Moreover, historical references are made to assert the presence of political continuity supposedly personified by contemporary Marxists. Termed as cultural Marxists, these individuals, following the failure of the political revolution, allegedly aspire to initiate a new cultural revolution with the objective of dismantling the gender binary and traditional family structures. Debates within this frame then expand to the “eternal themes” of Slovenian political space such as the existence of a deep state, supposedly represented by former communists seeking to perpetuate the exploitation of the nation. Some of these narrative frames are also echoed by Catholic and right-wing media Twitter accounts, which constitute an integral part of the fourth cluster.

*When you look at your crotch in the toilet, you know that there are only 2 possible sexes. Unless you're a cultural Marxist and you're wondering whether there are 59 or just 58. (@BernardBrscic)*

*Thought of the day:*

*We only have two types of leftists.*

*1. Leftists for personal gain #thieves, i.e. leftist criminals*

*2. Ideological leftists, who are mentally ill and should be sent to psychiatric hospitals. Poor things believe they have 70 genders. (@tomltoml)*

*The family is the fundamental unit of society, which cultural Marxists want to break up by any means necessary. They do this through LGBT propaganda, radical feminism and gender theory. Let us not allow them to destroy us. DEFEND FAMILY! (@Anti\_MigrantSLO)*

Community 5 adopts a specific discursive context centred on the discussion of language, linguistic norms, and rules. Here, "gender theory" enters the discourse through its intervention in language, particularly as it pertains to the actions taken at the Faculty of Arts. In this context, "gender theory" is framed as an intrusion into the domain of language, disrupting the otherwise value-neutral field of linguistics.

*Transgender at "Filofaks": Female student Janez Novak, you failed the exam given by female professor Božo Repe!<sup>9</sup> (...) "The introduction of the feminine grammatical gender for all genders has an exclusionary effect." (@PlusPortal)*

*It is neither about expertise nor common sense. The humanities are first and foremost an attitude towards the world and history. And this attitude absolutely excludes the ideological engineering of language. That is why every good humanist – like Ahačič – knows instinctively that such measures are incompatible with the spirit of Filofaks. (@cnfrmstA)*

Communities 6 and 7 consist of tweets that also appeared in the other communities. In Community 7, the distinctive element is the employment of humour and exaggeration to mock and undermine the legitimacy of "gender theory". These tweets attempt to highlight what they perceive as the absurdity of "gender theory", employing satirical elements to challenge its validity. The overall pattern observed across the tweets is based on the premise that "gender theory" exists, has a clearly defined nature, and operates according to the principles of conspiracy theories.

*A woman from Bohinj comes to Filofaks and says: "Oi! I wanna enrol, I'm gonna study, I'm gonna be a student, I'm gonna eat using vouchers".*

*\*Officer stares ahead for a few minutes, then performs a seppuku\* (@peterstroys)<sup>10</sup>*

*Our use of genders is also unknown in English. Student is student, he or she or it. Maybe they are preparing to study in foreign languages, habibi. (@DarkoErmenc)*

<sup>9</sup> Both names in the tweet (Janez Novak and Božo Repe) are male names.

<sup>10</sup> The humour of this tweet is very culturally conditioned. Bohinj is a Slovenian area where some women in the local dialect still use the masculine grammatical form for themselves, not the feminine. The original tweet: "pride Bohinjka na Filofaks, pa reče: 'žijo, se bi vpisu, bom študiru, bom študent, bom jedu na bone' \*referentka nekaj minut zabodeno strmi predse, nato izvede seppuku\*".

In our sample of the most retweeted tweets across all seven communities, we identified ten issues or problem statements that could be categorized into four groups. The first group consists of problem definitions based on the understanding of “gender theory” as a kind of conspiracy theory, which can result in the feminization of men, the endangerment of “our women”, or even in the extension of rights to new groups, including paedophiles and zoophiles.

The second group of problems is related to political tensions between the left and the right. The left, which supports LGBT and feminism, is defined as a problem, along with Islam and similar “dangers”, resulting from what is referred to on Twitter as “leftist-faggot delusions” or “degenerate leftism”.

The third set of problems is related to the relationship between freedom of speech and hate speech. The Twitter community argues that their freedom of speech is violated due to the persecution of their views on gender and that expressing the truth is now considered hate speech.

The final set of problems was specifically related to the intervention in the Slovenian language described above. This is believed to have led to the destruction of the Slovenian language and “ideological language engineering”.

Closely associated with problem definitions are the so-called active actors, those who contribute to the identified problems. Three groups of active actors were identified across all seven communities.

The first group of defined problems is linked to the LGBT community and radical feminists as active actors. George Soros is explicitly named as the one who allegedly provides financial support for the development of “gender theory”.

Actors from the second group are primarily defined by political tensions between the left and the right in Slovenia, including the ongoing debate about the legacy of Marxism and the role of different groups that fought in World War II. The actors causing the identified problems, therefore, include the “left government”, also referred to as the “primitives in power”, the second group consists of “cultural Marxists”, and the third group includes the “descendants of partisans”, who supposedly not only fought against Germany during World War II but also carried out a political revolution, enabling the victory of the communists.

The third group of active actors is closely linked to the events at the Faculty of Arts. The leadership of the Faculty of Arts is highlighted as a problem, specifically its dean, with a publicly known same-sex orientation.

Among the passive actors in all seven communities, i.e. those who are the victims of the identified problems, men (and their masculinity), women, family, school, and society in general are mentioned. Specifically related to the story from the Faculty of Arts, the Slovenian language and the employees of the Faculty who must adhere to new grammatical rules are identified as victims.

The analysed statements in the tweets contain fewer solutions than problems. Most of the tweets define the problem but do not offer explicit solutions, although the solution is implicitly contained in the problem definition, which simply suggests that the problem can be solved by “eliminating the problem”. For example, radical feminism, as an identified problem, can only be solved by “eliminating, abolishing” radical feminism. Similarly, the

prohibition of “gender theory”, especially in schools, is often suggested as a solution to the identified problems. The second group of solutions relates to calls for the protection of the family, traditional values, and morality. The final group of solutions specifically relates to the decision adopted at the Faculty of Arts, where the rejection of that decision and the correction of the mistake made are seen as the solution.

## 6 Discussion: Ideological Homogenization as a Consequence of “Gender Theory” Domination

Our findings provide empirical evidence supporting the analytical validity of the existing explanations of the characteristics and dynamics of “gender theory” and its role in public debates, but also point to the colonizing nature of “gender theory”, which moves away from its original thematization to new topics; by applying this flammable discursive framework, it is possible to transform nearly any topic into a first-class political ideological issue that divides the public.

Drawing upon the concept of the “symbolic glue” (Kováts & Põim, 2015), it has become evident that “gender theory” in Twitter debates functions as a unifying force, bringing together disparate actors who engage in separate debates in their respective communities, each with distinct thematic emphases, but with shared reference to “gender theory”. This shows how “gender theory” facilitates discursive cohesion among right-wing media, politicians, publicly known Slovenian right-wing figures (e.g. psychoanalysts), and, notably, numerous anonymous or publicly unknown users.

The inherent emptiness and malleability of “gender theory” as an “empty signifier” (Mayer & Sauer, 2017) in Twitter debates is observed in the fact that the identified communities introduced numerous new issues to the debate while discussing gender-sensitive language. This was made possible by the “glue” of the emptiness of the signifier “gender theory”. Specifically, our CFA shows that the communities focused on topics such as the integration of “gender theory” into the education system (Community 1), the defence of patriarchal roles (Community 2), the exposure of the dean’s homosexuality and allegations of childhood abuse (Community 3), the infiltration of “cultural Marxism” in educational institutions (Community 4), and the interpretation of the measure as an act of language destruction (Community 5).

The linguostylistic analysis also provides evidence supporting the understanding of “gender theory” as an “empty signifier”. This is exemplified by the identification of the conceptual opposition of “illness” vs. “health” in the most retweeted tweets. Within this framing, “gender theory” is represented as an “illness”, while “normalcy” is associated with “health”. Deviations from established grammatical norms are depicted as diseased and perverse, implying that society itself is afflicted by this condition, which is attributed to various elements such as radical feminism, the LGBT agenda, and cultural Marxism. The metaphorical use of “gender theory” demonstrates its emptiness, as it can encompass anything – in this instance, the contrast between “illness” and “health”.

The concept of the “ideological code” (Smith, 1999; Paternotte, 2023) becomes relevant when considering central users who adopt and propagate a signifier, unaware of its origin,

thereby enhancing its legitimacy and perceived accuracy in reflecting reality. In addition to right-wing media and politicians, active participation is observed from anonymous users who embrace the conceptual framework of “gender theory” as an unquestionable foundation for addressing gender-related issues. This empirical observation also aligns with Paternotte’s (2023) assumption that the control over “gender theory” has shifted away from its original creators, as it has now become widely disseminated among the general public.

Our empirical findings also demonstrate the role of “gender theory” in de-democratization (Lombardo et al., 2021) by limiting the boundaries of legitimate debate to the frame of “gender theory/ideology”, marginalizing and excluding “alternative” viewpoints and impeding democratic and pluralistic exchanges. Our analysis reveals the significant absence of any voices capable of offering alternative perspectives or framing the focal issue beyond the confines of “gender theory”. This indicates that the right-leaning network has effectively instrumentalized the Twitter debate, surpassing the *pro et contra* ideological polarization structure characterized by opposing viewpoints. Despite the identification of distinct communities, they ultimately stem from the same political network, creating an “echo chamber” where users through retweeting repeat ideologically similar posts, resulting in the marginalization and exclusion of alternative perspectives on the subject matter.

Put differently, we observe a departure from network “polarization”, which typically characterizes Twitter debates of this nature. Instead, our analysis reveals prevailing ideological homogenization across various communities. This phenomenon can be attributed to the colonizing nature of “gender theory”, which effectively restricts the space for the emergence of alternative viewpoints on a topic. Additionally, this trend may be attributed to the observed lower activity levels of left-leaning users on Twitter in Slovenia, which could be a result of the highly organized right-leaning political networks. Supporting this observation, Amon Prodnik (2016) highlights the existence of organized social media training within right-wing groups, a phenomenon that has not been similarly documented or reported for the left in Slovenia.

How “gender theory” contributes to de-democratization processes is also evident in the case of Community 3, employing psychoanalytic language to oppose the Faculty’s decision. Under the guise of scientific discourse, it sought to delegitimize the Faculty’s actions by pathologizing the dean, highlighting his homosexuality as the reason for the adoption of gender-sensitive measures and making allegations that he was abused sexually as a child. These attempts to frame disagreement as a personality pathology expose a troubling aspect of “gender theory” where seemingly rational (pseudo-psychoanalytic) discourse is employed to conceal malicious intentions.

Our study has limitations that should be considered when interpreting the findings. Firstly, our Twitter data collection was retrospective, covering a period in 2018 and collected in January 2023. Given the discontinuation of the Twitter API used, replicating this dataset is no longer feasible. Moreover, the dynamic nature of online digital environments implies that data collection at different times might yield varying results, even if the API were accessible.<sup>11</sup> Methodologically, our approach includes the use of the Louvain algo-

<sup>11</sup> The authors of this article maintain a database that is available to interested researchers, facilitating further study in this framework.

rithm for community detection, initially developed for undirected graphs, whereas our analysis involved a directed graph. Despite this, we selected the Louvain approach because of its success in some prior studies (Jesenšek et al. 2021; Smrdelj 2024) and Gephi version 0.9.2's capacity to handle directed and undirected graphs selected at initialization. However, using alternative algorithms or multiple applications of the Louvain algorithm could yield different outcomes. Nevertheless, the community structures we identified accurately reflect collaborations observed in the broader public. Lastly, our dataset covers the entire target population but is relatively small, limiting the generalizability of our findings to broader online digital publics.

## 7 Conclusion

In our study, we examined the Twitter discussion in Slovenia sparked by the decision of the Faculty of Arts, University of Ljubljana, to use gender-sensitive language in their internal regulation documents. Despite all the negative reactions, the Faculty has upheld the Senate's decision. Six years have passed since the resolution's adoption at the time of writing this text, and the Faculty continues to exercise this practice without reporting any issues regarding the comprehension of texts due to such language usage.

In our analysis of the Twitter responses to this decision, we identified distinct user communities that varied in size and retweeting activity. These communities were predominantly affiliated with right-leaning political networks. Through the linguostylistic analysis and CFA of the most retweeted tweets, we observed that the communities exhibited different thematic emphases regarding gender-sensitive language use and brought new topics into the discussion, but all these debates shared a common reference, which was the discursive concept of "gender theory".

To gain a more comprehensive understanding of the Twitter debate, future research should consider conducting longitudinal analyses and investigating the relationship between the actors involved in previous campaigns against equality and those adopting "gender theory". This would provide insights into coalition-building strategies and the persistence of "gender theory" in shaping public debates. However, future research should not only focus on longitudinal analysis, but it should also include advanced linguostylistic analysis. This includes the categorization of key words and phrases in the discourse and the mapping of these elements in the seven communities involved in the debate. Such an approach will allow a quantitative assessment of linguistic preferences and terminologies within each community and show how language reflects group ideologies and dynamics. In addition, analysing the most active users ("out-degree" centrality) will shed light on user engagement and the role individuals play in influencing the debate. Last but not least, a closer look at the edges' weight can provide a more comprehensive view of how discussions on "gender theory" evolve and spread over time.



## Acknowledgments

This research was conducted as part of the research programmes “Problems of Autonomy and Identities at the Time of Globalisation” (P6-0194, 2019–2027), “Slovene Language–Basic, Contrastive, and Applied Studies” (P6-0215, 2019–2027), and research project “Gender, Democracy and Neoconservative Anti-Gender Movement” (J5-50158, 2023–2026), all funded by the Slovenian Research and Innovation Agency.

## References

- Ahačič, K. (2018, May 26). Janez Novak, študentka slovenščine [Janez Novak, a female student of the Slovenian language]. *Delo, Sobotna priloga*. <https://www.delo.si/novice/slovenija/janez-novak-studentka-slovenscine/> Accessed: 16-07-2023.
- Al-Rawi, A. (2019). Twitter influentials and the networked publics' engagement with the Rohingya crisis in Arabic and English. In K. Smets, K. Leurs, M. Georgiou, S. Witteborn, & R. Gajjala (Eds.), *The SAGE Handbook of Media and Migration* (pp. 192–204). SAGE.
- Amon Prodnik, J. (2016). The instrumentalisation of politics and politicians-as-commodities: A qualitative analysis of Slovenian parties' understanding of political communication. *Annales, Series Historia et Sociologia*, 26(1), 145–158.
- Arguedas, A. R., Robertson, C. T., Fletcher, R. & Nielsen, R. K. (2022). *Echo Chambers, Filter Bubbles, and Polarisation: A Literature Review*. Reuters Institute University of Oxford.
- Blondel, V. D., Guillaume, J., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 10, P10008. <https://doi.org/10.48550/arXiv.0803.0476>
- Čechová, M., Krčmová, M., & Minářová, E. (2008). *Současná stylistika*. Praha: Nakladatelství Lidové noviny.
- Dombos, T., Krizsan, A., Verloo, M. & Zentai, V. (2012). *Critical Frame Analysis*. Center for Policy Studies, Central European University.
- Edenborg, E. (2023). Anti-Gender Politics as Discourse Coalitions: Russia's Domestic and International Promotion of »Traditional Values«. *Problems of Post-Communism*, 70(2), 175–184. <https://doi.org/10.1080/10758216.2021.1987269>
- Evkoski, Bojan, Mozetič, Igor, Ljubešić, Nikola & Petra Kralj Novak (2021): Community evolution in retweet networks. *PloS one*, (16)9, 1–19. <https://doi.org/10.1371/journal.pone.0256175>
- Fairclough, Norman. (1994). *Language and power*. New York: Longman.
- Ferra, I. & Nguyen, D. (2017). #Migrantcrisis: “tagging” the European migration crisis on Twitter. *Journal of Communication Management*, 21(4), 411–426. <https://www.emerald.com/insight/content/doi/10.1108/JCOM-02-2017-0026/full/html>
- Gorjanc, V. & Stabej, M. (2019). Slovenščina 2.0: »Jezik in družbeni spol« [Slovenian 2.0: “Language and Gender”]. *Slovenščina 2.0*, 7(2), i–iv. <https://doi.org/10.4312/slo2.0.2019.2.i-iv>

- Gorjanc, V., Golob Kalin, M., Kuhar, R., Močnik, R., Novak, B. A., Šorli, M., Štumberger, S. & Žele, A. (2019). Okrogla miza »Jezik in spol« [“Language and gender”: a round-table discussion]. *Slovenščina* 2.0, 7(2), 1–44. <https://doi.org/10.4312/slo2.0.2019.2.1-44>
- Graff, A. & Korolczuk, E. (2022). *Anti-Gender Politics in the Populist Moment*. Routledge. <https://doi.org/10.4324/9781003133520>
- Hansen, D. L., Shneiderman, B., Smith, M. A., & Himelboim, I. (2020a). Social network analysis: Measuring, mapping, and modeling collections of connections. In *Analyzing social media networks with NodeXL: insights from a connected world* (pp. 31–51). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-817756-3.00003-0>
- Hansen, D. L., Shneiderman, B., Smith, M. A., & Himelboim, I. (2020b). Twitter: Information flows, influencers, and organic communities. In *Analyzing social media networks with NodeXL: insights from a connected world* (pp. 161–178). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-817756-3.00011-X>
- Jeffries, L. & McIntyre, D. (2018). *Stylistics*. 11th printing. Cambridge University Press.
- Jesenšek, L., Verčič, D. & Kronegger, L. (2021). Struktura slovenske politične razprave na Twitterju ob volitvah v Evropski parlament [Structure of Slovenian political discussions on Twitter during European parliamentary elections]. *Teorija in praksa*, 58(1), 49–71. <https://www.doi.org/10.51936/tip.58.1.49-71>
- Kalin Golob, M. & Grizold, A. (2017). Slovenia’s state institutions in the grip of democratic political culture and the media: (the Telecom case). *World political science*, 13(2), 363–387. <https://doi.org/10.1515/wps-2017-0011>
- Kováts, E. & Pöim, M. (Eds.). (2015). *Gender as Symbolic Glue: The Position and Role of Conservative and Far Right Parties in the Anti-Gender Mobilization in Europe*. Foundation for European Progressive Studies, Friedrich Ebert Stiftung. <https://library.fes.de/pdf-files/bueros/budapest/11382.pdf> Accessed: 16-07-2023.
- Kubin, E. & von Sikorski, C. (2021). The Role of (Social) Media in Political Polarization: A Systematic Review. *Annals of the International Communication Association*, 45(3), 188–206. <https://doi.org/10.1080/23808985.2021.1976070>
- Kuhar, R. (2017). Changing gender several times a day: the anti-gender movement in Slovenia. In R. Kuhar & D. Paternotte (Eds.), *Anti-gender campaigns in Europe: mobilizing against equality* (pp. 215–232). Rowman & Littlefield.
- Kuhar, R. (2023). How the anti-gender movement contributed to marriage equality in Slovenia. In: E. Kováts, M. Laruelle & J. Chrobak (Eds.), *Culture wars in Europe: interdisciplinary perspectives on transformations in business-health relationships* (pp. 163–177). The George Washington University.
- Kuhar, R. & Paternotte, D. (Eds.). (2017). *Anti-gender campaigns in Europe: mobilizing against equality*. Rowman & Littlefield.
- Kuhar, R. & Pajnik, M. (2020). Populist mobilizations in re-traditionalized society: anti-gender campaigning in Slovenia. In G. Dietze & J. Roth (Eds.), *Right-wing populism and gender: European perspectives and beyond* (pp. 167–184). Transcript.

- Kuhar, R. & Antić Gaber, M. (2022). Women as a linguistic footnote: Equality fatigue in the debate on gender-sensitive language in Slovenia. *Gender and Language*, 16(3), 241–264. <https://doi.org/10.1558/genl.21680>
- Lombardo, E., Kantola, J., & Rubio-Marin, R. (2021). De-Democratization and Opposition to Gender Equality Politics in Europe. *Social Politics: International Studies in Gender, State & Society*, 28(3), 521–531. <https://doi.org/10.1093/SP/JXAB030>
- Martinsson, L. & Ericson, M. (2023). ‘Fight pandemics with protective masks or gender?’ Emerging collective identities and anti-gender movements on Twitter during the COVID-19 crisis in Sweden. *Feminist Media Studies*, 23(5), 2011–2027. <https://doi.org/10.1080/14680777.2021.2009898>
- Mayer, S. & Sauer, B. (2017). “Gender Ideology” in Austria: Coalitions around an empty signifier. In R. Kuhar & D. Paternotte (Eds.), *Anti-gender Campaigns in Europe: Mobilizing against equality* (pp. 23–40). Rowman & Littlefield.
- Metaxas, P., Mustafaraj, E., Wong, K., Zeng, L., O’Keefe, M. & Samantha, F. (2021). What Do Retweets Indicate? Results from User Survey and Meta-Review of Research. *Proceedings of the International AAAI Conference on Web and Social Media*, 9(1), 658–661. <https://doi.org/10.1609/icwsm.v9i1.14661>
- Mikić, J. & Kalin Golob, M. (2019). (Ne)strpno o spolu: Primerjava angleških in slovenskih argumentov proti uvajanju sprememb v seksistični rabi jezika [(In)tolerant on gender: comparison of English and Slovenian arguments against changing sexist language]. *Slovenščina 2.0*, 7(2), 75–117. <https://doi.org/10.4312/slo2.0.2019.2.75-117>
- Moon, S. J. & Hadley, P. (2014). Routinizing a New Technology in the Newsroom: Twitter as a News Source in Mainstream Media. *Journal of Broadcasting & Electronic Media*, 58(2), 289–305. <https://doi.org/10.1080/08838151.2014.906435>
- Ozuzden, O. & Korkut, U. (2020). Enmeshing the mundane and the political: Twitter, LGBTI+ outing and macro-political polarisation in Turkey. *Contemporary Politics*, 26(5), 493–511. <https://doi.org/10.1080/13569775.2020.1759883>
- Paternotte, D. (2023). Victor Frankenstein and its creature: the many lives of ‘gender ideology’. *International Review of Sociology*, 33(1), 80–104. <https://doi.org/10.1080/03906701.2023.2187833>
- Popič, D. & Gorjanc, V. (2022). Corpus-Linguistic Analysis of Speech Communities on Anti-Gender Discourse in Slovene. *Gender a výzkum / Gender and Research*, 23(2), 140–166. <https://doi.org/10.13060/gav.2022.020>
- Puente Núñez, S., Fernández Romero, D. & Maceiras, S. D. (2021). New Discourses of Masculinity in the Context of Online Misogyny in Spain: The Use of the “Feminazi” And “Gender Ideology” Concepts on Twitter. *Sociální studia / Social Studies*, 2, 49–66. <https://doi.org/10.5817/SOC2021-2-491>
- Smith, D. E. (1999). ‘Politically correct’: An organizer of public discourse. In *Writing the social: Critique, theory, and investigations* (pp. 172–194). University of Toronto Press.
- Smolej, M. (2022). Raba ženskega spola kot vključujočega – pro et contra [Using the Feminine Gender as Inclusive – Pro et Contra. The Case of the UL FF]. Primer UL FF. In: M. Antić

- Gaber (Eds.), *Samoumevnosti in nelagodja: Sistemsko urejanje enakosti spolov v akademskem prostoru* [Taken for Granted and Discomfort: Systemic Regulation of Gender Equality in Academia] (pp. 175–190). Založba Univerze v Ljubljani.
- Smrdelj, R. & Pajnik, M. (2022) Intersectional representation in online media discourse: reflecting anti-discrimination position in reporting on same-sex partnerships. *Gender, technology and development*, 26(3), 463–484. <https://doi.org/10.1080/09718524.2022.2144100>
- Smrdelj, R., Učakar, T. & Kuhar, R. (2021). Marginalizacija interseksijske perspektive: istospolna partnerstva in tujci na slovenskih novičarskih spletnih portalih [The Marginalisation of the Intersectional Perspective: Same-Sex Partnerships and Foreigners in Slovenian On-line News Portals]. *Javnost–The Public*, 28(supl.), S122–S140. <https://doi.org/10.1080/13183222.2021.2012945>
- Smrdelj, R. (2024). Communication Relations on Twitter during the Migrant “Crisis” in Slovenia. *Dve domovini/Two Homelands*, 59, 203–224. <https://doi.org/10.3986/2024.1.11>
- Šorli, M. (2019). Spol in jezik na robu pameti: med slovnico in ideologijo [At one’s wit’s end with language and gender: between grammar and ideology]. *Slovenščina* 2.0, 7(2), 45–74. <https://doi.org/10.4312/slo2.0.2019.2.45-74>
- Valicon (2020): Uporaba družbenih omrežij in storitev klepeta v Sloveniji 2018–2019 [The Use of Social Networks and Chat Services in Slovenia 2018–2019]. <https://www.valicon.net/sl/2020/01/uporaba-druzbenih-omrezij-in-storitev-klepeta-v-sloveniji-2018-2019/> Accessed: 09-01-2024.
- Van Bavel, J. J., Rathje, S., Harris, E., Robertson, C. & Sternisko, A. (2021). How Social Media Shapes Polarization. *Trends in Cognitive Sciences*, 25(11), 913–916. <https://doi.org/10.1016/j.tics.2021.07.013>
- Verloo, M. (2005). Mainstreaming gender equality in Europe: a critical frame analysis. *The Greek Review of Social Research*, 117(B), 11–34. <https://doi.org/10.12681/grsr.9555>
- Verloo, M. & Lombardo, E. (2007). Contested Gender Equality and Policy Variety in Europe: Introducing a Critical Frame Analysis Approach. In M. Verloo (Eds.), *Multiple Meanings of Gender Equality: A Critical Frame Analysis of Gender Policies in Europe* (pp. 21–49). CEU Press.
- Verweij, P. (2012). Twitter links between politicians and journalists. *Journalism Practice*, 6(5–6), 680–691. <https://doi.org/10.1080/17512786.2012.667272>
- Wallaschek, S., Kaushik, K., Verbalyte, M., Sojka, A., Sorci, G. Trenz, H. & Eig Müller, M. (2022). Same Same but Different? Gender Politics and (Trans-)National Value Contestation in Europe on Twitter. *Politics and Governance*, 10(1), 146–160. <https://doi.org/10.17645/pag.v10i1.4751>
- Wehling, E. (2020). Politics and framing. How language impacts political thought. In C. Cotter & D. Perrin (Eds.), *The Routledge Handbook of Language and Media* (pp. 136–150). Routledge.
- Yarchi, M., Baden, C. & Kligler-Vilenchik, N. (2021). Political Polarization on the Digital Sphere: A Cross-Platform, Over-Time Analysis of Interactional, Positional, and Affective Polarization on Social Media. *Political Communication*, 38(1–2), 98–139. <https://doi.org/10.1080/10584609.2020.1785067>

### Author's bio

Rok Smrdelj is Assistant Professor at the Department of Sociology, Faculty of Arts, University of Ljubljana (Slovenia). His research focuses on the sociology of media and communication, with a particular emphasis on the study of communication relations in the hybrid media sphere. He also explores structural changes in the public sphere, particularly in light of recent neoconservative mobilizations opposing equality policies in Europe and beyond. He received an award from the Slovenian Sociological Association for young, promising sociologists, and an award from the Faculty of Arts at the University of Ljubljana for the best doctoral dissertation in social sciences. Email: rok.smrdelj@ff.uni-lj.si

Roman Kuhar is Professor of Sociology at the Faculty of Arts, University of Ljubljana (Slovenia) and the former dean of the Faculty of Arts (2017–2021). His work has been published in numerous academic journals and he is the author of several books including (with A. Švab) *The Unbearable Comfort of Privacy* (Peace Institute, 2005), (with J. Takács) *Beyond the Pink Curtain: Everyday Life of LGBT People in Eastern Europe* (Peace Institute, 2007) and (with D. Paternotte) *Anti-gender Campaigns in Europe: Mobilizing against Equality* (Rowman & Littlefield, 2017). He is also associate editor at *Social Politics* (Oxford University Press) and the president of the Slovenian Sociological Association. Email: roman.kuhar@ff.uni-lj.si

Monika Kalin Golob is Professor of Linguistic and Media Stylistic at the Faculty of Social Sciences, University of Ljubljana (Slovenia) and the former dean of the Faculty of Social Sciences (2017–2021). Her research fields are language culture, stylistics of journalism, PR and advertising, language planning and language policy. Her bibliography is available on: [https://bib.cobiss.net/bibliographies/si/webBiblio/bib201\\_20230312\\_184242\\_15000.html](https://bib.cobiss.net/bibliographies/si/webBiblio/bib201_20230312_184242_15000.html)  
Email: monika.kalin-golob@fdv.uni-lj.si

---

KATA KNAUZ,\* ATTILA VARGA,\*\* ZSOLT SZABÓ\*\*\*  
✉ SÁRA BIGAZZI\*\*\*\*

Intersections. EEJSP  
11(1): 139–164.  
<https://doi.org/10.17358/ieejsp.v11i1.1227>  
<https://intersections.tk.hu>

## Mainstream Political Discourse on the Roma Minority in Hungary between 2010 and 2019

---

\* [\[knauzkatapszichologus@gmail.com\]](mailto:knauzkatapszichologus@gmail.com) (University of Pécs; Eötvös Loránd University, Budapest)  
\*\* [\[atis13@gmail.com\]](mailto:atis13@gmail.com) (Eötvös Loránd University, Budapest)  
\*\*\* [\[szabo.zsolt@ppk.elte.hu\]](mailto:szabo.zsolt@ppk.elte.hu) (Corvinus University of Budapest, Budapest, Hungary;  
Eötvös Loránd University, Budapest)  
\*\*\*\* [\[bigazzi.sara@pte.hu\]](mailto:bigazzi.sara@pte.hu) (University of Pécs)

### Abstract

This study examines how Hungarian politicians represented the Roma minority in public statements reported by the National News Agency from 2010 to 2019. Through an integrated theoretical framework, we analyzed communication styles to uncover not only what was said about the Roma but also the underlying motivations, proposed societal relationships, and social orders conveyed by these statements. Our methodology combined textual and thematic analyses to explore both the explicit content and the covert implications of these communications.

Our findings reveal a dual discourse from the governing party, Fidesz, which simultaneously emphasizes integration and inclusion to meet European expectations while exploiting anti-Roma sentiment and perceived victimhood, particularly during the refugee crisis, to appeal to domestic audiences. Fidesz employed a complex propaganda strategy, portraying the Roma as both subjects of integration and as scapegoats in the context of refugee-related threats. In contrast, far-right parties like Jobbik engaged in overtly racist propaganda, reinforcing negative stereotypes and solidifying in-group identity among their supporters. Liberal-left-centrist parties, meanwhile, were largely ineffective, often replicating Fidesz's portrayal of the Roma as helpless victims without providing substantive alternative narratives.

This analysis highlights the relational and deeply political nature of discourse, demonstrating how representations of the Roma are manipulated to serve broader political agendas. It also underscores the lack of critical engagement with structural inequalities, revealing how these narratives perpetuate existing power dynamics rather than addressing the systemic issues affecting the Roma community.

**Keywords:** political discourse; Roma; anti-Gypsyism; prejudice; inclusion; propaganda

## 1 Introduction

Since assuming power in 2010, Viktor Orbán's government in Hungary has embarked on a path of illiberal democracy characterized by the centralization of power and the undermining of established checks and balances (Enyedi & Krekó, 2018). This consolidation

---



extends to controlling political communication and public media (Krekó, 2022; Polyák, 2019). Despite Fidesz's significant parliamentary majority, the government cultivates an image of victimhood, resonating with the collective Hungarian psyche's perception of external and internal threats (Szabó & Csertő, 2023), with the Roma minority often positioned among these perceived threats.

On the European stage, Hungary presents a contrasting narrative. The European Union (EU) recognizes the Roma as Europe's largest ethnic minority and has actively worked to address their discrimination and social exclusion. EU initiatives emphasize socioeconomic inclusion, equality, and participation. However, the Roma community continues to face persistent challenges, with notable regional variation in prejudice and discrimination across EU Member States (Kende et al., 2021). Historical, cultural, and legal factors contribute to socially endorsed prejudice in Eastern European countries (Kende et al., 2017), in contrast to more veiled discrimination in Western European countries (Kende & McGarty, 2019).

Fidesz navigates these contradictory demands by consolidating power internally through narratives of victimhood and perceived threats while externally engaging in EU initiatives aimed at promoting minority inclusion and equality. The far-right opposition amplifies in-group victimhood narratives, often depicting Roma individuals negatively, deepening societal divides. Conversely, the liberal-left-centrist opposition struggles to articulate a distinct representation of the Roma that aligns with their values, hindered by entrenched societal biases (Bíró-Nagy et al., 2022).

Our research inductively examines Hungarian political discourse on the Roma, focusing on how this communication implicitly mediates political interests, the positions of the authors, and the proposed social relations. We examine how strategic communication and prejudice shape political agendas and national identity. A core premise of our study is that Hungarians' perceptions of the Roma are entwined with their self-identity construction (Bigazzi, 2009; Bigazzi et al., 2019). Utilizing moral typecasting theory (Gray & Wegner, 2009), which suggests that people categorize groups as either perpetrators or victims, we assess the implications of majority Hungarians viewing themselves through this lens. In particular, we explore how this self-perception influences Hungarian politicians' representation of the Roma, the nation's largest ethnic minority.

## 2 Theoretical Background

We construct our knowledge through communication, which is fundamentally mediated by representation (Duveen, 2002). Representation 'provides collectivities with intersubjectively shared means for understanding and communicating' (Duveen & Lloyd, 1990, p. 2). Moscovici (1961/1976) identified three distinct ways to represent and communicate a social object. Diffusion transmits content of general interest by blurring social differences among an audience with the aim not to influence openly but to inform, allowing opinions to be forged. Propagation transmits information via a more authoritarian relationship with a specific audience, integrating content into the latter's own worldview and values, assimilating some and neglecting others, thereby generating and strengthening attitudes. Propaganda has the aim of fostering group identity and identifying opponents and enemies.

These different communication processes shape the cognitive organization of knowledge of the audience differently (Buschini & Guillou, 2022). However, they also exhibit

different relationships between the source of information and the target group (audience), affording different degrees of autonomy in relation to the social object. Diffusion assumes a wide target group and encourages individual opinions, propagation frames content with implicit direction, and propaganda directly imposes relational attitudes on social objects. Power dynamics influence these processes, with dominant sources allowing for divergence and minority identities, often constructed on threat perception, driving toward cohesion and homogeneity (Deschamps, 1982). Propaganda, particularly, serves as a form of minority influence aimed at fostering societal change (Elchereth et al., 2011; Moscovici, 1961/1976).

Knowledge construction and representation are inherently relational, collaborative, and deeply political processes (Howarth, 2011). The source's political power dictates the domination of communication channels and the construction of identities and positioning. Communication about a social object shapes its representation, potential relations, and interactions. Additionally, it regulates and organizes the construction of the societal landscape by offering identity positions and placing the actors on the map of normative group relations.

The social representation of intergroup relations entails cognitive processes that shape the social order. Staerklé (2009; Staerklé et al., 2011, 2012) delineates four different mechanisms that regulate and organize social order: the moral order, rooted in shared values about bad and good; a free market system, based on meritocracy; social diversity, deriving from ethnic and cultural differentiation between groups; and structural inequality, centered on the management of inequality. The initial two mechanisms emphasize normative differentiation, while the latter are characterized by categorical differentiation.

Diffusion functions as a normalization process, providing reference frames (Sammut & Bauer, 2011) that support a broad, normatively aligned consensus. Propagation involves a majoritarian influence, leveraging a superordinate position to foster normative conformity. Conversely, propaganda serves as a minority influence strategy, challenging and proposing alternatives to the prevailing social order (Howarth, 2006). It frequently employs stigmatization or strong categorical differentiation to deepen social identity distinctions beyond the majority's normative parameters (Staerklé, 2015).

Relating to minorities involves navigating relational inequalities. Assimilation efforts (Gordon, 1964), often linked to prejudice, imply that those who are different lack normative qualities. Segregation creates physical and psychological distance from minorities, frequently through the relativization of differences, thereby strengthening categorization. Integration (Bourhis et al., 1997) focuses on the need for minorities to be part of the majority and, similarly to assimilation, implicitly expects action and responsibility from minorities (Bowskill et al., 2007). In contrast, inclusion represents a systemic approach that advocates for systemic changes to embrace diversity as a societal asset (Bigazzi et al., 2019). The fundamental distinction between integration and inclusion lies in their focus (Houtkamp, 2015). Relational concepts like inequalities often translate into policies designed to ameliorate or empower the target group to meet normative expectations without addressing relationships and their context. When differences are seen as the root of inequalities, policies targeting these groups aim to 'correct' these disparities. However, such approaches, despite their intentions, can inadvertently result in victim-blaming (Ryan, 1976).

The framework presented in this table encapsulates and typologizes the interplay between communication styles, the mental organization of knowledge, the relationships proposed by sources of information, the regulation behind social order, and the potential

societal relationships with minorities. In this research, we will explore how and what political actors communicate about the social object of the Roma, while also making assumptions about what this type of communication offers to its audience according to these theoretical dimensions (see Table 1).

**Table 1** Integrated theoretical framework of the typologies of political communication

Communication styles	Diffusion	Propagation	Propaganda	
Mental organization of knowledge	Opinion	Attitude	Stereotype	
Relationship with source	Normalization	Majoritarian influence	Minority Influence: aimed at societal change	
Social order	Free-market meritocracy	Moral order	Social diversity	Structural inequality
Differentiation	Normative	Normative	Categorical	
Relation to Minorities	Integration	Assimilation	Segregation	Inclusion

### 3 Hungarian Roma (Representations) Tied to Economic, Social, and Political Agendas

Messing and Bernath (2017), after analyzing Roma media representation in Hungary for 25 years, highlight the rarity of systematic changes in thematic framing, closely tied to shifts in political discourse. This suggests that the discourse reflects broader political and social positioning rather than the Roma's actual situation(s). We aim to outline the economic-political landscape since 1990, Roma policies, and their media representation.

The shift to a market economy in the 1990s adversely affected Roma people, with a significant proportion facing long-term unemployment and poverty (Bárány, 2001; Kertesi & Kézdi, 2011). The economic transition of the 1990s exacerbated structural inequalities, fueling interethnic tensions and a surge in public and political debates about national and local issues, including poverty, culture, crime, education, and Roma political participation (Feischmidt et al., 2013; Messing & Bernáth, 2017). The establishment of the Roma political representatives' system in 1993 coincided with the founding of Hungary's first extreme-right party, MIÉP (Party of Hungarian Justice and Life), whose main narrative portrayed the economic transition as a struggle between governing anti-Hungarian groups and national resistance (Kovács, 2013).

During the transition years, Hungary faced the dual challenges of democratization and Europeanization. The democratization process highlighted the insufficiency of merely establishing formal democratic institutions without their systemic integration into the socio-cultural context, leaving them ill-equipped to manage burgeoning informal political-business networks (Ágh, 2016; Herman, 2015). This situation led ordinary citizens to feel a loss of control over the transition, accompanied by concerns over the potentially high costs. Concurrently, Europeanization, marked by Hungary's 1994 EU membership application, required structural adjustments with a focus on ethnic minority issues, culminating in initiatives aimed at enhancing Roma social inclusion (European Council Framework

Convention, 1995). These efforts resulted in the development of policies and action plans to address the escalating issues of poverty, inequality in opportunities, and social exclusion faced by the Roma minority.

Despite these efforts, Roma employment plummeted in the early 2000s, highlighting persistent economic challenges and societal prejudice (Enyedi et al., 2004). The Roma issue became more prominent in the media, with a focus on social problems such as housing (Yuval-Davis et al., 2017). The rise of the radical right-wing party, Jobbik, with its anti-Roma stance, alongside Fidesz's ethnicization of national discourse, underscored growing societal divisions (Csigó & Merkovity, 2016; Kovács, 2013; Vidra & Fox, 2012). Particularly, after losing the election again in 2006, Fidesz intensified its minority influence propaganda, utilizing strong, 'empty populist' elements, such as 'direct reference to the will of the people, opportunism, the construction of a moral majority, and the promise of state defense against insecurity' (Csigó & Merkovity, 2016:300).

The global economic crisis in 2008 and the subsequent economic and moral collapse of the Hungarian government, famously illustrated by the socialist prime minister's admission of having lied to voters, amplified the crisis discourse and ethnic nationalism of Fidesz, as well as the overtly racist rhetoric of Jobbik. This shift led to a change in perspective towards the Roma minority, from viewing them as a national minority to ethnic 'othering.' The political and public discourse saw a significant escalation in overtly racist content, portraying the Roma as criminals (Vidra & Fox, 2014; Yuval-Davis et al., 2017). Jobbik's electoral success in 2009 and 2010 was closely linked to the emerging media attention on Roma-related issues. To gain legitimacy, Jobbik prioritized themes such as the 'Roma question' and 'Roma criminality' (Bíró-Nagy et al., 2012; Csiki, 2014; Feischmidt & Szombati, 2016).

During this period, violent incidents perceived as perpetrated by Roma individuals, such as the Olaszliszka lynching in 2006 and the Marion Cozma killing in Veszprém in 2009, heightened violence and prejudice against Roma. Threat marches by paramilitary organizations closely connected to Jobbik, such as the Hungarian Guard, were held, and racist attacks occurred that resulted in the death of six Roma individuals and several injuries, attributed to the extreme right 'gang of four,' associated with extreme right organizations (Csiki, 2014). A 2009 survey revealed widespread anti-Roma sentiment among Jobbik, Fidesz, and MSZP voters (Political Risk Index, 2010).

In 2010, following its electoral victory, Fidesz amended the Media Law, significantly restricting critical discourse and centralizing media control through the Media Council. This led to the consolidation of news content production via Hungary's national news agency, MTI, raising concerns about the autonomy and diversity of perspectives within the public media sphere. Critics were particularly worried about the potential impact on the independence and autonomy of Hungary's public media landscape (Hargitai, 2021).

Liberal-left-centrist parties were cornered in this communicative space: the emerging anti-Roma sentiment did not prompt these political parties to propose an alternative, more inclusionary perspective on the issue. The radical discourse was dominated by Jobbik, the only party that managed to present a stance independent of Fidesz, but this offered no solace to the Roma people. Jobbik's rhetoric focused on rejecting the elitism and corruption of mainstream parties alongside its anti-Roma and anti-Jewish positions (Batory, 2016).

The 2015 migration crisis further complicated Hungary's political landscape. The government's hardline stance on migration, juxtaposed with its claims of focusing on Roma integration, revealed a complex, often contradictory approach to minority issues (Tremlett & Messing, 2015).

Since 2010, the communication about Roma communities has been entangled in a politics of double discourse, where, on the one hand, their integration was promoted as part of the Europeanization process, and on the other, they remained internal, still subordinate ‘others’ to be dealt with (Kóczé & Rövid, 2019).

This historical context sets the stage for our research into the nuanced political discourse surrounding the Roma in Hungary since 2010. Our aim is to uncover the role of information sources of political actors, their broader interests within the political landscape, and how they influence the organization of social order and propose relationships with minorities.

## 4 Study

### 4.1. Research Question and Hypotheses

We sought to answer the following research questions:

How do political actors in Hungary communicate about Roma communities, and what underlying communication strategies and power dynamics are revealed through the analysis of overt and covert structures within this public discourse?

What does this communication reveal about their interests, roles, the relationships they propose to their audience, the social order they assume, and their approaches to minority groups?

Our study analyzed Roma-related communication by Hungarian politicians from 2010 to 2019, drawing on Messing and Bernáth’s (2017) observation that thematic shifts in the representation of Roma issues are rare and closely linked to broader political discourse changes. We expected to observe that two significant events during this period – the economic and moral collapse of the socialist government at the end of the 2000s and the refugee crisis in 2015 – would influence these discourses.

We explored the governing party’s efforts to navigate the expectations of both European and domestic audiences, hypothesizing that Fidesz politicians employ both propagation and propaganda to craft their messages, depending on the audience and the message’s specific objective. We anticipated that Jobbik would demonstrate a distinct rhetorical stance, differentiating itself from Fidesz’s ambiguous messages intended to placate both European partners and domestic audiences. Furthermore, we expected that other opposition parties, faced with strong anti-Roma sentiments among the electorate and a political space dominated by Fidesz and Jobbik’s narratives, which starkly contrast their core values, would likely opt for silence. Finally, we also expected that the migration crisis would alter the political discourse related to the Roma communities, with Fidesz using the Roma as a justification for anti-refugee policies.

### 4.2 Data

In a pilot project in 2018, we collected data from popular news websites, such as index.hu, hvg.hu, 24.hu, 444.hu, 888.hu, and kurucinfo.hu. However, we soon realized that most relevant articles identified MTI as their source. Despite the controversies surrounding the independence and autonomy of MTI (see Hargitai, 2021, for example), in terms of Hungarian

politicians' communication about the 'Gypsy' and 'Roma' communities, there was a nearly complete overlap between the news provided by MTI and the news appearing on these portals.

Thus, we decided to systematically gather the relevant statements from MTI using the keywords 'Gypsy' and 'Roma' for the period between 2010 and 2019. Our inclusion criteria included statements that referred to the Hungarian Roma community and were made by Hungarian politicians from either the government parties (Fidesz, KDNP) or opposition parties that received at least 1 percent of votes in elections. This threshold is meaningful as parties failing to achieve it must repay campaign support. We also included communications from official Roma minority representatives but excluded statements from non-affiliated public figures or those broadly concerning Roma at the European level (i.e., not specifically Hungarian Roma).

This process yielded 1,829 statements with 290,891 total words and 1,931,088 characters, averaging 190.25 words per statement. Most statements originated from Fidesz (44.72 percent) and Jobbik (20.40 percent), with a noticeable decline in statements since the 2015 refugee crisis (see Table 2).

Table 2 Breakdown of Statements by Year and Source

Source	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	Total
Governing Parties and Government											
Fidesz-KDNP	84	120	138	144	79	45	67	60	47	33	817
Government	5	3	3	5	9	19	32	26	14	11	127
Liberal-Left-Centrist Opposition Parties											
MSZP	43	36	35	19	21	6	9	11	2	6	188
LMP	24	23	24	8	4	13	1	5	3	2	107
DC	NA	NA	7	12	12	8	2	NA	4	6	51
Together	NA	NA	NA	11	11	4	1	NA	NA	3	30
Liberal party	NA	NA	NA	NA	1	2	4	2	NA	2	11
Momentum	NA	NA	NA	NA	NA	NA	NA	NA	NA	1	1
Radical Right-Wing Parties											
Jobbik	56	80	74	59	37	19	14	11	5	18	373
OHM	NA	NA	NA	NA	NA	NA	NA	NA	2	14	16
Other											
Independent	13	1	3	3	3	1	1	2	NA	6	33
Roma Self-go.	12	22	4	24	9	1	1	2	NA	NA	75
Grand Total	237	285	288	285	186	118	132	119	77	101	1829

Note: DC = Democratic Coalition. OHM = Our Homeland Movement. Roma Self Go. = Roma Self Government. NA = Not Available.



### 4.3 Method

In order to uncover the complex interplay between communication styles, the role of the source, and how the audience is expected to relate to it, as well as the proposed social order and relationships with minorities as outlined in our theoretical framework, we employed two complementary inductive methods of text analysis. These methods were specifically chosen for their ability to reveal both the overt and covert structures within political discourse, aligning with our aim to explore how political actors communicate about the Roma in Hungary.

The first method, a form of automated multidimensional text analysis using IraMuTeQ, leverages graph theory and the co-occurrence of words to identify the main clusters within public discourse. This approach aligns with our theoretical interest in communication styles and the organization of knowledge, as it allows us to map the ‘visible’ and ‘said’ aspects of the discourse—essentially, the dominant narratives and how they are structured within the broader social order. By automatically detecting patterns across large datasets, this method helps us identify prevalent themes and communication strategies employed by political actors.

The second method, a manual bottom-up thematic analysis, complements the automated approach by delving deeper into the broader textual and social context of the discourse. This method is crucial for interpreting the meanings behind statements and understanding the deeper structures and patterns related to how minorities are represented and the underlying intentions of the information sources. It is particularly well-suited to uncovering implicit relations and power dynamics that may not be immediately visible through automated analysis alone.

Together, these methods offer a comprehensive view of the data from two distinct but interconnected perspectives: the first method provides an overarching map of the discourse, while the second allows us to interpret deeper, more nuanced patterns. This dual approach is particularly effective in uncovering the layered meanings and social implications of political communication concerning Roma communities—specifically, how and why political actors discuss these communities and what underlying information and patterns are revealed according to the assumptions in our theoretical framework.

The automated textual analysis of the 1,829 statements involved several steps. Initially, we used DeepL to translate the statements into English and removed non-alphanumeric tokens for consistency. We then utilized IraMuTeQ software, which is well-suited for handling theoretical frameworks emphasizing communication and the social construction of knowledge (Sarrica et al., 2016). IraMuTeQ enables the quantitative analysis of a corpus through various tools: lexical analysis, similarity analysis, descending hierarchical analysis (DHA), and correspondence factor analysis (CFA). The establishment of the corpus included morphological normalization, converting words to their base forms for lemmatization, and was conducted using the software’s English dictionary (Pola et al., 2015). Lexical analysis identifies and reformats text units into initial context units (ICUs) and elementary context units (ECUs). It performs word frequency identification and lemmatization, simplifying words to their root forms for clustering. This process standardizes verbs to their indicative tense, reduces plurals to singular forms, and simplifies variants of lexical roots to capture essential lexicons for statistical analysis (Sarrica et al., 2016). Similarity analysis employs graph theory to identify co-occurrences between words and their hierarchical

significance or structure (Marchand & Ratinaud, 2012; Pereira, 2001). DHA facilitates the classification of text segments based on lemmatized word frequency, maximizing similarity within classes and differences between them. The chi-square ( $\chi^2$ ) test measures relationships between words by comparing observed and expected distributions. Higher  $\chi^2$  values indicate the greater specificity of a term within a word class, resulting in a partition of classes aiming for homogeneity and differentiation. CFA, following DHA, plots words and variables linked to each DHA cluster on a Cartesian plane, highlighting significant words based on p-values. Words are positioned according to their frequency or  $\chi^2$  values in their class.

For the manual thematic analysis, we adopted an approach based on the framework established by Braun and Clark (2013). We developed a three-level coding system consisting of codes (defined as the most fundamental units of analysis, representing parts of the data deemed relevant for the analysis), themes (defined as patterns within the data, indicating a meaningful combination of codes), and overarching themes (defined as a higher level of abstraction, embodying broad, more general themes that emerge from the organization and integration of more specific themes in the data). Initially, the first author created a coding system through deep immersion, adopting a bottom-up approach. This in-depth examination enabled the first author to identify keywords, subsequently establishing patterns among these codes and organizing them into themes and overarching themes. The process was iterative, aimed at challenging emerging patterns and identifying cases that did not conform. Once the coding system was established, the first and second authors of the paper coded the data independently. The statements were divided into two corpora, with each corpus fully coded by one coder and a randomly chosen 10 percent of the statements coded by the other. With the goal of preserving the data's natural representation and recognizing interconnected themes across varying levels of abstraction, our coding system was non-exclusive, allowing for a single statement to be assigned multiple codes or themes. The interrater reliability, assessed using Cohen's Kappa, demonstrated an average agreement of 98.46 percent, reinforcing the credibility and consistency of our coding process.

## 4.4 Results

### 4.4.1 Automatized Textual Analysis

The lexical examination of 1,829 statements produced 9,833 text segments (TS), with a total of 347,504 occurrences. The Hapax index stood at 1.02, showing that 35.19 percent of the words (lemmas) were used only once: 3,547 out of 10,080 forms. We focused our analysis on adjectives, adverbs, unrecognized forms, common names, and verbs, reducing the dataset to 7,359 forms. In this refined dataset, 2,501 hapaxes (forms used only once) represented 0.72 percent of total occurrences and 33.99 percent of forms and were excluded from further analysis.

The first step was a word cloud analysis which, after removing the search words 'Roma' and 'Gypsy', as well as 'Hungarian,' revealed that public discourse on Roma during this period mainly highlighted political actors like government (N = 1882), Europe (N = 1214), and minister (N = 1375, including references to Zoltán Balog, then Minister of Human



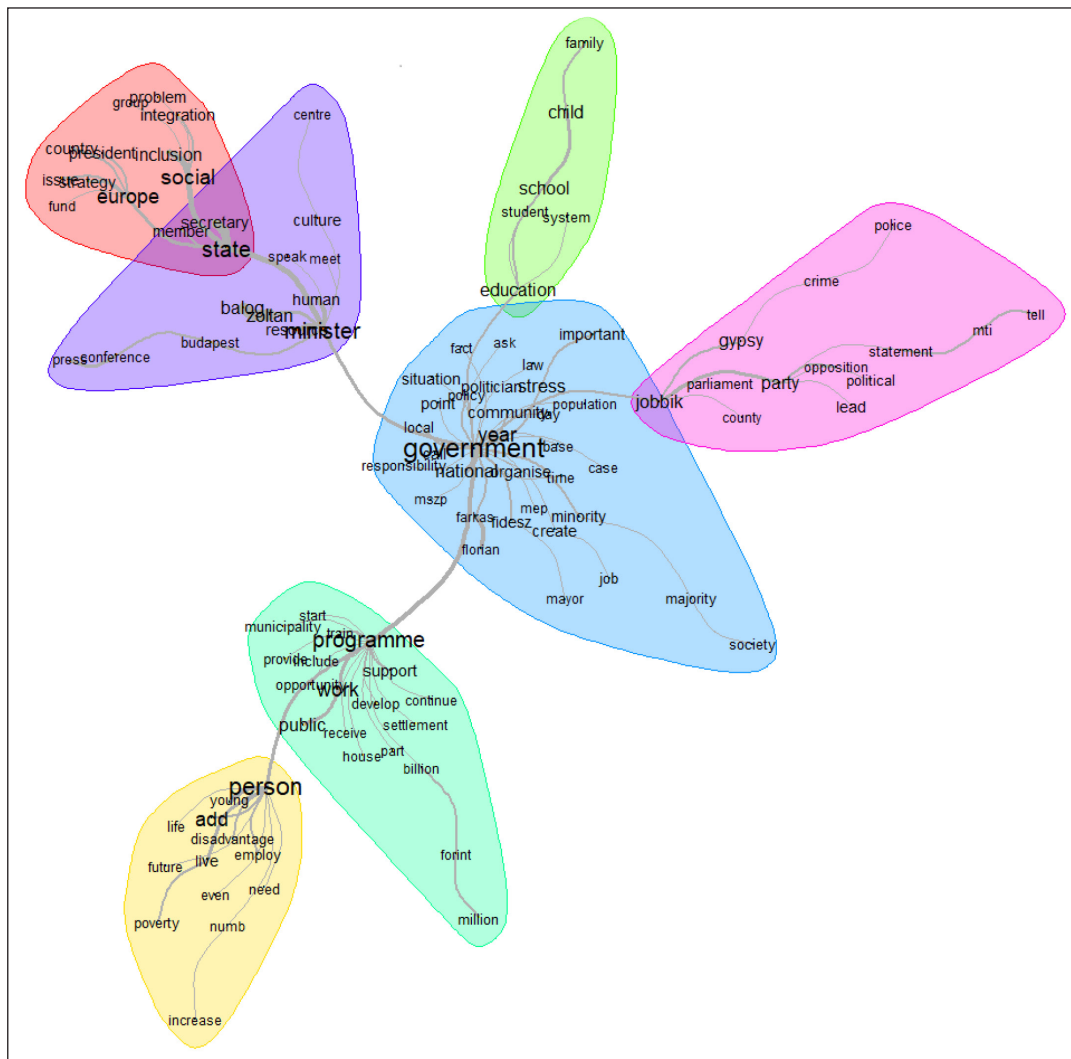


Figure 1 Similarity Analysis: Co-Occurrences of Words with Frequencies Greater Than 200

Third, we conducted a descending hierarchical analysis with an 85.09 percent success rate in classifying text segments, indicating a satisfactory index (Camargo & Justo, 2013). The analysis yielded five classes, hierarchically organized into three clusters. Figure 2 provides a summary of these results, showing the classes, clusters, and their relationships. Each class in the figure is highlighted by its most significant words, selected based on their  $p$  values ( $< .0001$ ) and the highest chi-square values. Cluster A, Integration, includes Class 1 (Education) and Subcluster C (European Discourse) with Class 2 (EU Dialogue) and Class 3 (Culture). Cluster B, National Discourse, comprises Class 4 (Local Conflicts) and Class 5 (Parties).

These findings reveal a dual discourse on the Roma community between 2010 and 2019. The first discourse (Cluster B), rooted in the national context, involves politicians from various parties who leverage the issue for debates (campaign, elect), assign blame (racism, hate), and express opposition (condemnation). This discourse focuses more on political positioning than on constructive discussions about the Roma issue (coexistence). The modalities of our variable source (authors) saturating in this class are the different opposition parties and Roma representatives. The local context, marked by significant intergroup conflict (e.g., Miskolc, Gyöngyöspata), is characterized by demonstrations and marches that often require police intervention to maintain safety, order, and prevent crime. The modalities of our variable source saturating in this class are Jobbik and Our Homeland.

The other discourse (Classes 1, 2, 3) focuses on achieving inclusion, emphasizing education (school, child, disadvantage, vocational) and identity politics rooted in cultural specifics (music, community, value, history). This discourse is framed within the broader European context and political landscape (inclusion, strategy, Europe, European Framework, administration, policy, affair). The modalities of the variable source saturating in these three classes are governmental and FIDESZ representatives.

This duality illustrates the complex and fluid representation of communication about Hungarian Roma, which is influenced by political agendas and societal tensions. In the national context, the Roma become a tool for political objectives, perpetuating stereotypes and exacerbating social divisions. In the political news related to the European context, they are the focus of integration strategies and social welfare initiatives.

In Figure 2, we can observe the time variable saturating in each class. The two national classes are predominantly characterized by news before the refugee crisis, while the two European classes and the Inclusion class are more representative of communication in the period following the refugee crisis.

The automated text analysis of news related to the Roma community reveals key patterns in how political actors and parties frame their discourses. A focus on the most frequent words indicates that the discourse predominantly centers on integration, emphasizing strategies to reduce poverty, prevent crime, and address social disadvantages. Education is identified as the primary means to achieve these goals.

Further analysis of the nodes and edges in the similarity analysis shows that this discourse is centrally organized around the government, which is depicted as pivotal in shaping and directing public conversation about Roma issues. This central governmental community is connected to several other nodes the Minister of Human Resources, Zoltán Balog, who communicates about integration; European strategies of inclusion and funding; and the political opposition, Jobbik, which uniquely criminalizes the Roma, presenting a divergent political strategy in Hungary. Education also emerges as a key community, with its perceived role in the integration process being underscored. The Programs and Initiatives discourse is primarily focused on individuals living in poverty and facing disadvantages. However, this narrative also includes the allocation of significant financial resources—billions of forints—toward programs designed to meet these individuals' needs.

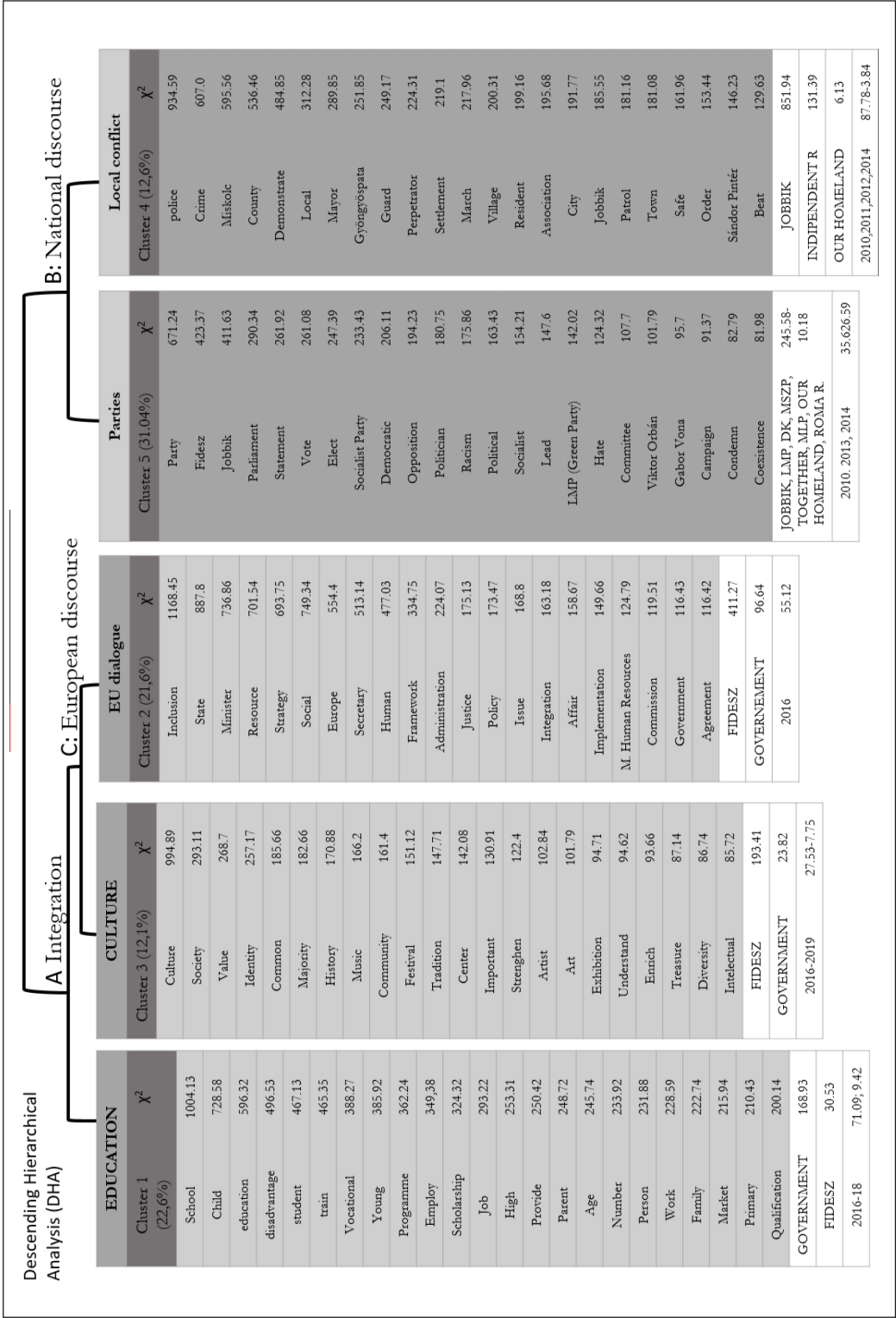


Figure 2 Descending Hierarchical Analysis Results: Cluster and Class Structure and Hierarchical Relationships (Numbers in Parentheses Indicate the Percentage of the Total Text Segments Included in Each Class)



The clustering analysis reveals a notable shift in political communication before and after the refugee crisis. Before the crisis, political discourse about the Roma in MTI news was dominated by national debates between parties, reflecting a competitive political environment with multiple voices vying for influence. After the crisis, the discourse shifted to emphasize EU frames, education, and identity politics, with Fidesz and government actors dominating. The content became less diverse, with Jobbik's narrative of 'gypsy crime' emerging as the only significant dissenting voice. This shift from diffusion to propagation—marked by fewer, more controlled messages—reflects the government's consolidation of power and control over the narrative. By this time, the Roma discourse had come to focus on cultural identity politics rather than addressing the structural inequalities involved in Roma integration.

The choice of communication strategy, whether diffusion or propagation, serves to shape the proposed social order. Before the refugee crisis, the more open diffusion approach allowed for a plurality of voices, suggesting a competitive political environment. After the crisis, the shift to propagation indicates a consolidation of power by the ruling party, reflecting a more hierarchical social order. The government's focus on cultural identity over structural inequalities reinforces the superficial inclusion of the Roma while leaving deeper issues of economic and social inequality unaddressed. Rather than fully aligning with European values, Fidesz engages selectively, meeting EU expectations on surface-level issues like cultural identity while resisting substantial reforms that could address economic and social inequalities. This strategy minimizes opposition and maintains the ruling party's dominance in the political landscape. The shift from a broader, inclusive discourse to a controlled narrative underscores the government's intent to focus on political expediency rather than genuine change for the Roma community, perpetuating a social order that superficially includes the Roma but fails to address the deeper inequalities that sustain their marginalization.

#### 4.4.2. Thematic Analysis

In the manual thematic analysis, since codes were non-exclusive, we assigned 5,570 codes across 1,829 statements, with an average of 3.05 codes per statement (range: 1 to 10). We identified seven overarching themes through this analysis that encapsulate a wide range of themes and codes. For brevity, we will focus on these overarching themes in this section, with a detailed list available in the appendix. The overarching themes are presented in order of their frequency, from the most to the least common.

The 'Helping the Roma' overarching theme, representing 33.68 percent of codes (N = 1876), includes strategies for Roma integration and descriptions of EU-funded projects, portraying Roma as beneficiaries of support. Fidesz dominates this discourse, accounting for 64.57 percent of the mentions, followed by liberal-left-centrist parties (12.53 percent), Jobbik (2.82 percent), and Our Homeland Movement (0.05 percent). Fidesz's statements often present the Roma as disadvantaged and in need of governmental support, typically conveyed in a patronizing manner, highlighting Fidesz and/or the government's efforts to improve conditions for the Roma. Liberal-left-centrist parties have attempted to propose alternative approaches, yet they have still portrayed the Roma as helpless. Far-right parties

suggest helping in a distinctly patronizing way, such as proposing boarding schools for Roma children to improve their situation by removing them from their familial environments. The frequency of this theme declined after peaking in 2013. A sample quote for this overarching theme comes from an event at a summer camp for disadvantaged children, at which the Minister assured the Roma children that they could count on the Government for social inclusion in school and job creation, stating, 'We want you to succeed in life.' (Zoltán Balog, then Minister of Human Resources, 01.07.2016)

The 'Roma as Victims' overarching theme, comprising 19.07 percent (N = 1062) of the codes, emerged as a significant narrative. This involves portraying the Roma community as facing adversities like poverty, segregation, and racism. This theme included historical references to the Holocaust and hate crimes. Fidesz was the leading contributor (49.52 percent), followed by liberal-left-centrist parties (37.74 percent), Jobbik (4.23 percent), and Our Homeland Movement (0.09 percent). A peak in this narrative occurred in 2013. Additionally, the analysis shows that although the liberal-left-centrist opposition has significantly contributed to voicing this narrative of victimhood, their emphasis on it has diminished over the years. A sample quote for this overarching theme comes from Zoltán Balog, who drew attention to the need to take child poverty seriously, as it narrows the world of those affected and makes both the environment and those living in it more distrustful. 'Child poverty is also closely linked to parents' educational attainment, social status, inadequate housing conditions, territorial disadvantages, and ethnicity,' he said. Referring to the latter, he indicated that 'children from Roma families are more affected by poverty' (Zoltán Balog, then Minister of Human Resources, 19.09.2013)

Approximately 17.59 percent (N = 980) of the codes were related to the overarching theme of 'Interethnic Relations,' involving discussion of Roma integration and relations with the majority. Topics included cohabitation issues, often framing Roma as a societal challenge and emphasizing their perceived integration failure. Fidesz contributed the most codes (52.14 percent), followed by Jobbik (17.85 percent) and liberal-left-centrist parties (14.18 percent). A peak in Fidesz-related codes occurred in 2013, with a decline afterward. The discourse, while not overtly hostile, often involved a paternalistic tone, suggesting that Roma issues could only be resolved by higher authorities, thus marginalizing Roma agency and ignoring structural inequalities. The liberal-left-centrist opposition's statements were subtler, advocating peaceful coexistence and integration. We provide two sample statements for this overarching theme, one from the government and one from the opposition.

The government considers the future and integration of the Roma population in Hungary a matter of national strategy, as stated in the opening speech of the government during a parliamentary debate entitled 'Chances for the Integration of the Roma Population in Hungary' held on Tuesday. 'The improvement of the situation of this mass of 700-750 thousand people, within which currently 80 per cent are unemployed, is a guarantee for the growth of the Hungarian economy or the sustainability of the pension system,' Zoltán Balog indicated. The Minister referred to the establishment of a system of officials within the government responsible for dealing with issues associated with the Roma population as an important government measure. Zoltán Balog expressed that 'The Roma issue is much more important than [can be] entrust[ed...] solely to the Roma (...) but it is much more important [to deal with it] than to exclude the Roma from it. Quoted from Zoltán Balog, then Minister of Human Resources (Hungarian News Agency, 2012, November 20).

Viktor Szigetvári stated that the leadership of Together would hold a meeting in Miskolc in August to find the best solutions to local issues of coexistence. He indicated that they would also consult with local Roma organizations and other parties. He added that Together's local members had presented a fair and pro-integration program during the parliamentary campaign and that they would continue to engage in fair politics; for that reason, they supported them becoming local councilors. Quoted from Viktor Szigetvári, then leader of the leftist-liberal party, Together (Hungarian News Agency, 2014, July 23).

Approximately 12.69 percent (N = 707) of the codes are categorized under the overarching theme of Rightist rhetoric, highlighting deviancy, 'Gypsy crime,' and the perceived threat from the Roma community, including concerns about their growing population. This theme encompasses discussions on punishments, restrictions, opposition to positive discrimination, and accusations of Roma selling their votes. It critiques political correctness regarding the Roma issue. Overall, these codes each contribute to a broader right-leaning political discourse, casting the Roma community as perpetrators and the majority population as victims. Jobbik leads this narrative (74.82 percent), followed by Fidesz (11.73 percent), Our Homeland Movement (6.93 percent), and liberal-left-centrist parties (3 percent). A sample quote for this overarching theme is:

The real problem,' according to Zsolt Egyed, is that the majority of the Roma population in Hungary 'live in 19th-century conditions with a 17th-century mentality... They have settled for the fact that the best way to make a living is by having children, exploiting the social welfare system, and engaging in crime'. Quoted from Zsolt Egyed, MP of the radical right-wing party Jobbik (Hungarian News Agency, 2012, November 20).<sup>2</sup>

Nearly twelve percent (11.74 percent; N = 654) of the codes focused on the 'Responsibility' of various actors for the situation of the Roma community. Most statements (57.49 percent) from opposition parties highlight the government's accountability. Jobbik's responsibility for the situation of the Roma is noted in 11.79 percent of the codes, while the former socialist government's accountability is mentioned in 18.2 percent of the statements. Additionally, 10.86 percent of the codes directly attribute disadvantages faced by the Roma to the community itself. A sample quote for this overarching theme is:

Tímea Szabó spoke about the dire educational situation and low level of education among most of the Roma population in Hungary. She emphasized that it is unacceptable that the governments after the regime change were unable or unwilling to address this problem and that a third generation is growing up without going to school. Quoted from Tímea Szabó, then MP of the green party LMP (Hungarian News Agency, 2012, November 13).

The 'Positive Aspects' overarching theme emphasized the Roma community's positive characteristics, including culture, music, heroes, and their value to society, making up 4.43 percent (N = 247) of all codes. Most of these codes were attributed to Fidesz politicians (69.23 percent), with additional contributions from government representatives (15.79 percent), followed by liberal-left-centrist opposition parties (7.92 percent), and last, by Jobbik (2.02 percent). However, these statements often transmitted mixed messages, recognizing Roma culture but also perpetuating stereotypes and reinforcing a sense of otherness. A sample quote for this overarching theme is provided in a statement of the Minister of Human Resources, Zoltán Balog, who highlighted that: „ *The festival sends a message to*

*those living in deep poverty, who need something to hold onto the most, that they have a culture worth living for, worth getting to know, and worth embracing.*”—quoted from Zoltán Balog, then Minister of Human Resources (Hungarian News Agency, 2014, May 11).

Finally, the ‘Anchoring Migrants to Roma’ overarching theme emerged, linking Roma and refugee issues. Though less frequent, making up only 0.79 percent (N = 44) of all codes, it significantly influenced the portrayal of the Roma in political discourse, especially during 2015–16. Fidesz politicians were responsible for 81.82 percent of these statements, with Jobbik and government spokespersons each contributing 6.82 percent. The government cited overwhelmed social systems and Roma issues as reasons to reject refugees, framing both Roma and refugees as non-European outsiders. A sample quote for this overarching theme is:

Hungary also has no right to comment on the experiments of other countries in this regard. In connection with this, Viktor Orbán referred to Hungary’s historical circumstances, stating that it is inherent for Hungary to coexist with a few hundred thousand Roma people, which, according to him, should not be objected to by anyone in any way. However, he also emphasized, ‘we do not impose the demand on anyone (...) that they too should live together with a large number of Roma minority. Quoted from Viktor Orbán, Prime Minister of Hungary (Hungarian News Agency, 2015, July 09)

Summarizing the manual thematic analysis of Hungarian political discourse on the Roma community reveals a narrative characterized by paternalism, victimization, and exclusion. Fidesz’s portrayal of the Roma as dependent on government aid reinforces a hierarchical dynamic, diminishing Roma agency while upholding state authority. Although the discourse framing Roma as victims was initially prominent across political parties, it has declined, particularly among the opposition. Right-wing rhetoric, dominated by Jobbik, perpetuates harmful stereotypes, portraying the Roma as a societal threat and justifying exclusionary policies. This narrative underscores the deep-seated influence of far-right ideologies in Hungarian politics. Additionally, the shifting of responsibility for Roma issues—whether onto the government, previous administrations, or the Roma themselves—reflects a broader reluctance to address systemic challenges. Overall, Hungarian political discourse continues to reinforce power structures and marginalize the Roma, with no political actors systematically addressing the root causes of inequality, instead perpetuating dependency and exclusion.

This analysis identifies a propaganda strategy within the right-wing rhetoric, where portraying the Roma as a societal threat instills fear and reinforces negative stereotypes, solidifying the political base by presenting the Roma as a common enemy. Similarly, Fidesz’s linking of Roma issues to the refugee crisis reflects a propaganda approach. By framing both groups as threats to Hungarian society and as non-European outsiders, Fidesz justifies exclusionary policies, consolidates nationalistic sentiments, and constructs minority influence in Europe concerning refugee policies.

Fidesz’s portrayal of Roma as disadvantaged and in need of state intervention aligns with a propagation strategy. This approach reinforces the party’s image as a benevolent protector while also subtly maintaining a hierarchical relationship where the Roma are dependent on state assistance. The emphasis on government accountability and, to a lesser extent, on the Roma’s responsibility fits into this propagation strategy, aiming to shape public opinion about who is at fault for the Roma’s situation.

Discussions around Roma integration, cultural identity, and the positive contributions of the Roma reflect a diffusion strategy. These narratives, although less prominent, attempt to disseminate more balanced or diverse perspectives, often highlighting cultural values but still tinged with paternalism. These discussions are less about directing public opinion and more about introducing alternative viewpoints into the broader discourse, though these viewpoints are often overshadowed by dominant narratives.

## 5 Discussion

Through a combination of textual and thematic analysis of political discourse on the Roma from 2010 to 2019, the study aims to understand how Hungarian politicians represent the Roma community in their statements. Grounded in social representation theory, our study suggests that representation is inherently relational, collaborative, and deeply political (Howart, 2011). This approach offers insights not only into the representation of the Roma but also into the position of the communicators, their proposals for social order, and their relationship with both the imagined audience and minority representatives within the political community.

Social representation theory identifies three ways to represent a social object: diffusion, propagation, and propaganda. Diffusion, which aims to inform and allow audiences to form their own opinions, was observed in the diverse voices within the political arena, particularly before the refugee crisis. This style of communication allowed for a plurality of opinions, where various political actors, including opposition and minority representatives, presented their viewpoints. This open communication style reflects a meritocratic aspect of social order, where different ideas, even if only superficially, compete for acceptance in the public sphere.

Propagation seeks to integrate content into the audience's worldview, reinforcing certain attitudes while downplaying others. Fidesz's use of propagation is evident in the consistent framing of the Roma as dependent on government support. This approach reinforces a moral order in which the state is seen as the protector and provider, shaping societal norms and expectations regarding minority integration.

Propaganda, by contrast, is used to create stereotypes that strengthen the identity of the perceiving group while identifying adversaries. The use of propaganda, particularly by Jobbik and Fidesz during the refugee crisis, illustrates a shift towards a more explicitly authoritarian communication style. This style fortifies in-group identity by delineating the Roma, and by extension, refugees, as out-groups, thereby steering the social order towards exclusion rather than inclusion. This approach also serves as a minority influence in the broader European political arena.

The Hungarian scenario during the specific period presented a unique case of power dynamics. Fidesz, the ruling party, wielded constitutional power yet depicted its supporters and the broader Hungarian population as besieged. This narrative resonated with a historical sense of victimhood deeply embedded in Hungarian society (Szabó, 2020). Such positioning of victimhood typically frames others as antagonists (Gary & Wegner, 2009), and studies indicate that Hungarians harboring exclusive victim beliefs displayed anti-refugee, anti-EU, and anti-Roma sentiments, driven by perceived threats from these groups (Szabó et al., 2020; Szabó & Csertő, 2023; Szabó et al., 2023).



Fidesz navigated domestic anti-Roma sentiments and EU directives on Roma integration with an ambiguous approach, representing Roma as helpless victims and as dependents needing authority intervention. This strategy, blending propagation and propaganda, entrenched stereotypes without addressing systemic inequalities or fostering empowerment. It essentialized inequalities to the Roma minority deemed in need of 'improvement' ('felzárkóztatás,' meaning 'improving them,' often misused in Hungarian political discourse as a synonym for inclusion), a dynamic Ryan (1976) termed victim blaming. Integration efforts seemed to focus on elevating minority members to meet social, educational, and labor norms rather than adopting an inclusive strategy that addresses the relations between privileged and disadvantaged groups and aims at systemic change. This double-edged communication by Fidesz was augmented by voicing powerful, overtly right-wing statements that revealed Fidesz's true position, especially during the refugee crisis when the Roma were used as a proxy to highlight the perceived threats posed by refugees. Such analogies, effective in shaping public perceptions (Ghilani et al., 2017), highlight Fidesz's deliberate depiction of the Roma as a symbolic threat to Hungarians. Fidesz argued that integrating the Roma was already a Herculean effort, suggesting that existing resources should be prioritized for the Roma, with whom 'we have to live,' as notably stated by Viktor Orbán. They implied that Hungary has inherited the challenge of dealing with 'a few hundred thousand Roma.' In essence, Fidesz, initially responding to the economic downturn in 2008 with a diverse audience in mind, employed an ambiguous system of representation. Following the refugee crisis, it purposefully recast the Roma as a figurative peril to the Hungarian populace.

Jobbik, emerging in response to Hungary's economic and moral crisis, embraced a narrative that vilified the Roma as threats to societal stability, employing propaganda to solidify its base. Jobbik positioned itself as the party for non-Roma individuals negatively affected by the systemic changes, criticizing the flawed democratization and Europeanization processes. For Jobbik, the Roma community represented a clear threat, central to mobilizing its voters. The representational void left by Jobbik was filled by Our Homeland Movement, which exceeded expectations in the 2022 elections, indicating a receptive audience for its overtly racist rhetoric.

Liberal-left-centrist parties struggled to differentiate their representation of the Roma from that of Fidesz or Jobbik. Similar to Fidesz, they often portrayed the Roma as victims in need of assistance but did so in a patronizing manner, failing to establish genuine partnerships.

The propagation of opinions and attitudes that view Roma primarily through a lens of dependency or criminality suggests a stereotypical organization of knowledge within the audience. This framing simplifies complex socioeconomic issues into binary good-versus-bad narratives, often neglecting the systemic factors contributing to the marginalization of the Roma. Our manual analysis highlighted how stereotypes are reinforced through both overt propaganda by right-wing parties and more subtle propagation methods by Fidesz. These stereotypes not only affect the social image of the Roma but also constrain their opportunities, reflecting a social order that favors segregation and discrimination over integration.

Fidesz's dominant role in the discourse normalizes its perspectives as the majority's views, thereby shaping national understanding of Roma issues. This majoritarian influence is a powerful tool in shaping public opinion and guiding legislative approaches to minority integration.



Before the refugee crisis, the diffusion of diverse voices suggested an attempt at meritocracy where political ideas could compete. However, the dominant propagation and propaganda strategies employed by Fidesz and Jobbik have since shifted the discourse toward a moral order in which the state and certain political entities dictate social norms and values. This top-down approach emphasizes a hierarchical social structure where the government positions itself as the moral and practical authority on minority issues. The communication strategies reveal a tension between a professed commitment to social diversity and the perpetuation of structural inequality. While political discourse occasionally highlights the positive aspects of Roma culture, suggesting a recognition of social diversity and promoting identity politics, the overwhelming portrayal of Roma as dependent or criminal reinforces structural inequalities. This inconsistency indicates a superficial acknowledgment of diversity that fails to challenge or change underlying structural inequalities.

The government's communication, particularly through propagation, nominally supports integration by focusing on education and social programs. However, the paternalistic and often stereotypical portrayal of the Roma, coupled with the endorsement of voices like Jobbik's, who advocated for Roma children to be placed in boarding schools, leans toward segregation under the guise of integration. This political approach not only undermines the agency of the Roma community but also entrenches their status as outsiders within their own country. The strategies oscillate between assimilation—whereby Roma are expected to conform to mainstream societal norms without adequate support or recognition of existing barriers—and inclusion, which remains more rhetorical than practical. True inclusion, which addresses structural inequalities, remains elusive in the political rhetoric analyzed here.

Our study faces limitations, notably regarding the source of the analyzed political discourse. The revision of the Media Law in 2010 sparked controversy over the independence and autonomy of the Hungarian News Agency (MTI). Despite our initial pilot project showing that MTI is a prevalent source for most relevant articles in news portals, an alternative interpretation of our findings persists. There is a possibility that liberal-left-centrist political parties have developed a distinct representation of the Roma community but struggled to disseminate their perspectives to a broader audience. While our study cannot conclusively verify this hypothesis, it is crucial to recognize that reliance on MTI as our data source may have introduced bias into our data and, consequently, affected our interpretations.

## 6 Conclusion

Our comprehensive analysis of political discourse related to the Roma community in Hungary from 2010 to 2019 reveals a predominantly categorical differentiation in the portrayal of the Roma. Political entities have leveraged communication strategies that oscillate between diffusion, propagation, and propaganda, significantly influencing public perceptions and societal integration of the Roma.

Jobbik's approach, characterized by explicit propaganda, underscores a clear demarcation between the Roma and the majority, often involving portraying the Roma as a societal threat. This strategy possibly entrenches negative stereotypes and also exacerbates social division, reflecting a broader European right-wing trend of employing fear-based politics to mobilize electoral support.

Conversely, Fidesz has employed a more nuanced strategy that blends propagation with elements of propaganda. While ostensibly advocating for Roma integration and welfare, Fidesz's discourse often maintains a paternalistic tone, suggesting the dependency of the Roma on government initiatives. This mixed messaging creates a facade of inclusivity while subtly reinforcing existing power hierarchies and social inequalities. Such strategies, though aligning superficially with European values on minority integration, actually serve to maintain Fidesz's political dominance by controlling the narrative around minority issues.

Liberal-left-centrist parties, while occasionally critiquing governmental approaches, have largely failed to propose effective alternative strategies or engage substantively with the Roma issue. Their limited impact on the discourse around Roma integration highlights a significant gap in advocacy and policymaking from parties traditionally viewed as advocates for minority rights.

In conclusion, while the political discourse superficially addresses the integration and welfare of the Roma, the underlying strategies reveal a complex landscape where power dynamics, European policy influences, and national politics converge. This convergence shapes not only public perceptions but also the socio-political realities facing the Roma community in Hungary. The need for a genuine and inclusive approach to Roma integration—one that goes beyond political rhetoric and addresses systemic inequalities—is more critical than ever.

## References

- Ágh, A. (2016). The Decline of Democracy in East-Central Europe: Hungary as the Worst-Case Scenario. *Problems of Post-Communism*, 63(5-6), 277–287. <https://doi.org/10.1080/10758216.2015.1113383>
- Bárány, Z. (2001). The socio-economic impact of regime change in Eastern Europe: Gypsy marginality in the 1990s. *East European Politics and Societies*, 15(1), 64–113.
- Batory, A. (2016) Populists in government? Hungary's "system of national cooperation", *Democratization*, 23(2), 283–303. <https://doi.org/10.1080/13510347.2015.1076214>
- Bigazzi, S. (2009). *Gypsy representation–Gypsy identity*. Unpublished doctoral dissertation. Università di Roma "La Sapienza", Italy.
- Bigazzi, S., Serdült, S., & Bokrétás, I. (2019). Empowerment of intergroup harmony and equity. D. Christie & N. Balvin (Eds.) *Children and Peace, From Research to Action*. Peace Psychology Series (pp. 119–137). Springer.
- Bíró-Nagy, A, Boros, T., & Varga, A. (2012). *Right-wing extremism in Hungary*. Friedrich-Ebert-Stiftung. International policy analysis. [https://www.academia.edu/download/49112482/Rightwing\\_extremism\\_in\\_Hungary.pdf](https://www.academia.edu/download/49112482/Rightwing_extremism_in_Hungary.pdf)
- Bíró-Nagy, A., Szászi, Á., & Varga, A. (2022). *Divided Hungary: Political polarization of the Hungarian society*. Friedrich-Ebert-Stiftung. Retrieved from: [https://www.policysolutions.hu/userfiles/Policy\\_Solutions\\_Divided\\_Hungary\\_2022\\_EN.pdf](https://www.policysolutions.hu/userfiles/Policy_Solutions_Divided_Hungary_2022_EN.pdf)
- Bourhis, R. Y., Moise, L. C., Perreault, S., & Senecal, S. (1997). Towards an Interactive Acculturation Model: A Social Psychological Approach. *International Journal of Psychology*, 32(6), 369–386. <https://doi.org/10.1080/002075997400629>

- Bowskill, M., Lyons, E., & Coyle, A. (2007). The rhetoric of acculturation: When integration means assimilation. *British Journal of Social Psychology*, 46(4), 793-813. <https://doi.org/10.1348/014466607X182771>
- Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners* (1st ed.). SAGE Publications Ltd. ISBN: 9781847875815
- Buschini, F., & Guillou, E. (2022). Diffusion, propagation, propaganda: And then came effusion. A new mode of communication for social representations. *Papers on Social Representations*, 31(2), 11.1–11.21. <http://psr.iscte-iul.pt/index.php/PSR/article/view/613/515>
- Camargo, B.V., & Justo, A.M. (2013). IRAMUTEQ: um software gratuito para análise de dados textuais. *Temas em psicologia*, 21(2), 513–518. <https://www.redalyc.org/pdf/5137/513751532016.pdf>
- Csiki, T. (2014). *Radical Right-Wing Attempts and Government Responses to Constructing and Securitized a 'Roma Issue' in Hungary, 2006–2010*. Thesis submitted in CEU nationalism studies program. [https://www.etd.ceu.edu/2014/csiki\\_tamas.pdf](https://www.etd.ceu.edu/2014/csiki_tamas.pdf)
- Csigó, P., & Merkovity, N. (2016). Hungary: Home of empty populism. In Aalberg, T., Esser, F., Reinemann, C., Strömbäck, J., & Vreese, C. H. (Eds.) *Populist political communication in Europe* (pp. 299-310). Routledge.
- Deschamps, J.-C. (1982). Social identity and relations of power between groups. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 85–98). Cambridge University Press.
- Duveen G.(2002) Construction, belief, doubt. In: Apostolidis G., Duveen N., & Kalampalikis, N. (Eds.). *Représentations et croyances. Psychologie & Société*, 5, 139–156. Paris: Laboratoire Européen de Psychologie Sociale
- Duveen, G. and Lloyd, B. (1990). *Social Representations and the Development of Knowledge*. Cambridge University Press.
- Elchereth, G., Doise, W., & Reicher, S. (2011). On the knowledge of politics and the politics of knowledge: How a social representations approach helps us rethink the subject of political psychology. *Political Psychology*, 32(5), 729–758. <https://doi.org/10.1111/j.1467-9221.2011.00834.x>
- Enyedi, Zs., Fábíán, Z., & Sik, E. (2004). Is Prejudice Growing in Hungary? In Kolosi, T., Vukovich, Gy., & Tóth, I. Gy. (Eds.) *Social Report 2004* (pp. 363–385.). Budapest: TÁRKI. <https://tarki.hu/sites/default/files/adatbank-h/kutjel/pdf/a737.pdf>
- Enyedi, Z., & Krekó, P. (2018). Explaining Eastern Europe: Orbán's laboratory of illiberalism. *Journal of Democracy*, 29(3), 39–51. <https://doi.org/10.1353/jod.2018.0043>
- Feischmidt, M., & Szombati, K. (2016). Understanding the rise of the far right from a local perspective: Structural and cultural conditions of ethno-traditionalist inclusion and racial exclusion in rural Hungary. *Identities*, 24(3), 313–331. <https://doi.org/10.1080/1070289X.2016.1142445>
- Feischmidt, M., Szombati, K., & Szuhay, P. (2013). Collective criminalization of the Roma in Central and Eastern Europe: Social causes, circumstances and consequences. In S. Body-Gendrot, M. Hough, K. Kerecsi, R. Lévy, & S. Snacken (Eds.). *The Routledge Handbook of European Criminology* (pp. 168–187). Routledge. <https://doi.org/10.4324/9780203083505.ch1>

- Flament, C. (1962). L'analyse de similitude. *Cahiers du Centre de Recherche Opérationnelle*, 4, 63–97
- Ghilani, D., Luminet, O., Erb, H.-P., Flassbeck, C., Rosoux, V., Tames, I., & Klein, O. (2017). Looking forward to the past: An interdisciplinary discussion on the use of historical analogies and their effects. *Memory Studies*, 10(3), 274–285. <https://doi.org/10.1177/1750698017701609>
- Gordon, M. (1964). *Assimilation in American Life: The Role of Race, Religion, and National Origins*. Oxford University Press
- Gray, K., & Wegner, D.M. (2009). Moral typecasting: Divergent perceptions of moral agents and moral patients. *Journal of Personality and Social Psychology*, 96(3), 505–520. <https://doi.org/10.1037/a0013748>
- Hargitai, H. (2021). Content and framing in radio news bulletins in urban and rural Hungary. *European Journal of Communication*, 36(2), 125–147. <https://doi.org/10.1177/0267323120966839>
- Herman, L.E. (2015). Re-evaluating the Post-Communist Success Story: Party Elite Loyalty, Citizen Mobilization and the Erosion of Hungarian Democracy. *European Political Science Review* 7(1), 251–284. <https://doi.org/10.1017/S1755773914000472>
- Houtkamp, C. (2015). Beyond assimilation and integration: The shift to ‘national’ and ‘transnational’ inclusion. *Acta Universitatis Sapientiae, European and Regional Studies*, 8(1), 73–87. <https://doi.org/10.1515/auseur-2015-0014>
- Howarth, C. (2006). A social representation is not a quiet thing: Exploring the critical potential of social representations theory. *British Journal of Social Psychology*, 45(1), 65–86. <https://doi.org/10.1348/014466605X43777>
- Howarth, C. (2011). Representations, identity and resistance in communication. In: D. Hook, B. Franks, & M.W. Bauer (Eds.), *The Social Psychology of Communication* (pp. 153–168). Palgrave MacMillan. [https://doi.org/10.1057/9780230297616\\_8](https://doi.org/10.1057/9780230297616_8)
- Kende, A., Hadarics M., & Lasticova, B. (2017) Anti-Roma attitudes as expressions of dominant social norms in Eastern Europe. *International Journal of Intercultural Relations*, 60(2), 12–27. <https://doi.org/10.1016/j.ijintrel.2017.06.002>
- Kende, A., Hadarics, M., Bigazzi, S., Boza, M., Kunst, J. R., Lantos, N. A., Lášticová, B., Minescu, A., Pivetti, M., & Urbiola, A. (2021). The last acceptable prejudice in Europe? Anti-Gypsyism as the obstacle to Roma inclusion. *Group Processes & Intergroup Relations*, 24(3), 388–410. <https://doi.org/10.1177/1368430220907701>
- Kende, A., & McGarty, C. (2019). A model for predicting prejudice and stigma expression by understanding target perceptions: The effects of visibility, politicization, responsibility, and entitativity. *European Journal of Social Psychology*, 49(5), 839–856. <https://doi.org/10.1002/ejsp.2550>
- Kertesi, G., & Kézdi, G. (2011). Roma employment in Hungary after the post-communist transition. *Economics of Transition*, 19(3), 563–610. <https://doi.org/10.1111/j.1468-0351.2011.00410.x>
- Kóczé, A., & Rövid, M. (2019). The Europeanisation of racial neoliberalism: The case of “Roma” and “Refugees”. In: I. Cortés Gómez, & M. End (Eds.), *Dimensions of Antigypsyism in Europe* (pp. 107–123). European Network Against Racism and the Central Council of German Sinti and Roma in Brussels. [https://www.enar-eu.org/wp-content/uploads/20116\\_book\\_roma\\_final.pdf](https://www.enar-eu.org/wp-content/uploads/20116_book_roma_final.pdf)

- Kovács, A. (2013). The post-communist extreme right: The Jobbik party in Hungary. In: R. Wodak, M. Khosravini, & B. Mral (Eds.), *Right-wing Populism in Europe: Politics and Discourse* (pp. 223–234). Bloomsbury Academic. <https://library.oapen.org/bitstream/handle/20.500.12657/58743/1/9781780933436.pdf>
- Krekó, P. (2022). The birth of an illiberal information autocracy in Europe: A case study on Hungary. *Journal of Illiberalism Studies*, 2(1), 55–72. <https://doi.org/10.53483/WCJW3538>
- Marchand, P., & Ratinaud, P. (2012). L'analyse de similitude appliquée aux corpus textuels: les primaires socialistes pour l'élection présidentielle française (septembre-octobre 2011). *JADT 2012: Actes des 11ème Journées internationales d'Analyse statistique des Données Textuelles*. 687–699.
- Messing, V., & Bernáth, G. (2017). Disempowered by the media: Causes and consequences of the lack of media voice of Roma communities. *Identities*, 24(6), 650–667. <https://doi.org/10.1080/1070289X.2017.1380264>
- Moscovici, S. (1961/1976). *La psychanalyse: Son image et son public*. Paris: Presses Universitaires de France (2nd ed., 1976).
- Pereira, C. (2001). Análise de dados qualitativos aplicados às representações sociais. *Psicologia*, 15(1), 177–204. <https://doi.org/10.17575/rpsicol.v15i1.495>
- Pola, L. G., Sarrica, M., & Contarello, A. (2015). Imprenditori di identità a Palazzo Marino. Cittadinanza e valori nei discorsi di insediamento dei Sindaci di Milano dal dopoguerra a oggi. *Psicologia Sociale*, 10(3), 223–256. <https://doi.org/10.1482/81370>
- Polyák, G. (2019). Media in Hungary: Three pillars of an illiberal democracy. In: (...) Połńska, E., & Beckett, C. (Eds.), *Public service broadcasting and media systems in troubled European democracies* (pp. 279–303). Springer.
- Ryan, W. (1976). *Blaming the victim*. Vintage Books.
- Sammut, G., & Bauer, M.W. (2011). Social influence: Modes and modalities. In: D. Hook, B. Franks, & M.W. Bauer (Eds.), *The Social Psychology of Communication* (pp. 87–106). Palgrave Macmillan. [https://doi.org/10.1057/9780230297616\\_5](https://doi.org/10.1057/9780230297616_5)
- Sarrica, M., Mingo, I., Mazzara, B., & Leone, G. (2016). The effects of lemmatization on textual analysis conducted with IRaMuTeQ: results in comparison. *JADT 2016: 13ème Journées Internationales d'Analyse Statistique des Données Textuelles*. 1–10.
- Staerklé, C. (2009). Policy attitudes, ideological values and social representations. *Social and Personality Psychology Compass*, 3(6), 1096–1112. <https://doi.org/10.1111/j.1751-9004.2009.00237.x>
- Staerklé, C. (2015). Social order and political legitimacy. In: G. Sammut, E. Andreouli, G. Gaskell, & J. Valsiner (Eds.), *The Cambridge Handbook of Social Representations* (pp. 280–294). Cambridge University Press.
- Staerklé, C., Clémence, A., & Spini, D. (2011). Social representations: A normative and dynamic intergroup approach. *Political Psychology*, 32(5), 759–768. <https://doi.org/10.1111/j.1467-9221.2011.00839.x>
- Staerklé, C., Likki, T., & Scheidegger, R. (2012). A normative approach to welfare attitudes. In S. Svallfors (Ed.), *Contested Welfare States: Welfare Attitudes in Europe and Beyond* (pp. 81–118). Stanford University Press.



- Szabó, Z. P. (2020). Studied and understudied collective victim beliefs. What have we learned so far and what's ahead? In: J.R. Vollhardt (Ed.), *The Social Psychology of Collective Victimhood* (pp. 163–185). Oxford University Press.
- Szabó, Z.P., & Csertő, I. (2023). The associations between exclusive collective victim beliefs and negative attitudes toward the EU. *International Journal of Intercultural Relations*, 96, 101851. <https://doi.org/10.1016/j.ijintrel.2023.101851>
- Szabó, Z.P., Mészáros, N.Z., Kővágó, P., & Fülöp, É. (2023). Az Összehasonlító Kollektív Áldozati Vélekedés Kérdőív magyar változatának pszichometriai jellemzői [Psychometric properties and validity of the Hungarian version of the Collective Comparative Victim Beliefs Questionnaire]. *Magyar Pszichológiai Szemle*, 78(1), 81–112. <https://doi.org/10.1556/0016.2023.00018>
- Szabó, Z.P., Vollhardt, J.R., & Mészáros, N.Z. (2020). Through the lens of history: The effects of beliefs about historical victimization on response to refugees. *International Journal of Intercultural Relations*, 74, 94–114. <https://doi.org/10.1016/j.ijintrel.2019.10.009>
- Tremlett, A.C., & Messing, V. (2015, August 4). *Hungary's future: anti-immigration, anti-multiculturalism and anti-Roma?* Open Democracy <https://www.opendemocracy.net/en/can-europe-make-it/hungarys-future-antiimmigration-antimulticulturalism-and-antiro/> Accessed: 10-03-2024
- Vergès, P., & Bouriche, B. (2001). L'analyse des données par les graphes de similitude. *Sciences Humaines*, 1–90. <https://www.scienceshumaines.com/sites/default/files/2025-01/Bouriche.pdf>
- Vidra, Z., & Fox, J. (2012). *The Rise of the Extreme Right in Hungary and the Roma Question: The radicalization of media discourse*. CEU Policy Brief. <https://www.cps.ceu.edu/publications/policy-briefs/rise-of-extreme-right-hungary-and-roma-question-2012> Accessed: 10-03-2024
- Vidra Z., & Fox, J. (2014) Mainstreaming of Racist Anti-Roma Discourses in the Media in Hungary. *Journal of Immigrant & Refugee Studies*, 12(4), 437–455. <https://doi.org/10.1080/15562948.2014.914265>
- Yuval-Davis, N., Varjú, V., Tervonen, M.M., Hakim, J. & Fathi, M. (2017). Press discourses on Roma in the UK, Finland and Hungary. *Ethnic and Racial Studies*, 40(7), 1151–1169. <https://doi.org/10.1080/01419870.2017.1267379>

## Data Sources

- Hungarian News Agency. (2012, November 13). Accessed via PET registration. Available at: <https://mti.hu> (Accessed: 15 October, 2020).
- Hungarian News Agency. (2012, November 20). Accessed via PET registration. Available at: <https://mti.hu> (Accessed: 15 October, 2020).
- Hungarian News Agency. (2012, November 20). Accessed via PET registration. Available at: <https://mti.hu> (Accessed: 15 October, 2020).
- Hungarian News Agency. (2014, May 11). Accessed via PET registration. Available at <https://mti.hu/> (Accessed: March 10, 2021)



Hungarian News Agency. (2015, July 09). Article accessed via PET registration. Available at <https://mti.hu/> (Accessed: June 10, 2021)

Hungarian News Agency. (2014, July 23). Accessed via PET registration. Available at <https://mti.hu/> (Accessed: March 10, 2021)

## Appendix

List of the Overarching Themes (I., ..., VII.), Themes (1., ..., 8.), and Codes (a., ..., e.) in the Thematic Analysis.

Overarching theme I. Roma as victims. Themes: 1. Poverty, segregation. 2. Anti-Roma attitudes, racism: a. Inciting hatred, b. Demonstrations and marches. 3. Remembrance of Anti-Roma violence: a. Remembrance of the Holocaust, b. Remembrance of the murders of the Roma, c. Hate crimes, d. Historical disadvantages.

Overarching theme II. Rightist rhetoric. Themes: 1. Deviance: a. Gypsy crime, b. Inability to integrate, c. Socialization, d. Threatened public safety, e. Unwillingness to work. 2. Perceptions of threats: a. Anti-Hungarian racism, b. Tension, c. Roma as a danger to Hungarians, d. Increasing Roma population, e. Childbearing for economic sustenance. 3. Punishment, restrictions. 4. Against positive discrimination. 5. Questioning political correctness. 6. Cheap Roma votes.

Overarching theme III. Interethnic relations. Themes: 1. Roma integration. 2. Gypsy-Hungarian relation: a. Problems of cohabitation, b. Societal issues, c. Gypsy question, Roma issue, d. Promotion of peaceful cohabitation.

Overarching theme IV. Responsibility. Themes: 1. Accountability of the government: a. Government is not strict enough, b. Government betrays the Roma, c. Corruption. 2. Jobbik's accountability. 3. Roma people's responsibility.

Overarching theme V. Helping the Roma. Themes: 1. Lessons. 2. Cooperation: a. For the Roma, b. With the Roma. 3. Education, schooling. 4. Equity: a. Equality is already present, b. Development of the social system, c. Legal equality. 5. Help from the government. 6. Strengthening Roma identity. 7. Decreasing unemployment. 8. Unemployment as a problem.

Overarching theme VI. Positive aspects. Themes: 1. Culture and music. 2. The Roma heroes. 3. Roma people as a resource/value: a. Roma belong to the nation.

Overarching theme VII. Anchoring to migrants. Themes: 1. Supporting the Roma instead of migrants: a. Roma compared to the migrants.

- 
- \* [Szalay\_Aron@student.ceu.edu] (Research Center for Computational Social Science,  
Faculty of Social Sciences, ELTE Eötvös Loránd University, Budapest, Hungary;  
Department of Sociology and Social Anthropology, Central European University, Vienna, Austria)
- \*\* [zsofia.rakovics@tatk.elte.hu] (Research Center for Computational Social Science,  
Faculty of Social Sciences, ELTE Eötvös Loránd University, Budapest, Hungary;  
Doctoral School of Sociology, ELTE Eötvös Loránd University, Budapest, Hungary)

### Abstract

This study investigates the prevalent use of fear as a political tool by Hungarian Prime Minister Viktor Orbán. Through critical discourse analysis and automated text analytics, we analyze the State of the Nation addresses from 2010 to 2022, identifying recurring enemy images such as ‘Brussels’, ‘migrants’, ‘George Soros’, ‘Ferenc Gyurcsány’ and ‘opposition.’ National Consultations are also examined to understand fear-inducing strategies. Our findings reveal a consistent pattern of fearmongering, with enemy images aligning with key events and political contexts.

By employing both qualitative and quantitative methods, this research deepens our understanding of fear’s role in shaping public perceptions and political dynamics. It sheds light on the strategic use of fear in Hungarian politics and its impact on democratic processes. Furthermore, the study highlights the implications for the construction of illiberal systems in Central and Eastern Europe.

The analysis of Orbán’s discourse provides valuable insights into the manipulation of public sentiment and the consolidation of power through fear tactics. This study contributes to the broader discourse on the intersections of fear and politics, providing insight into fearmongering through official speeches.

**Keywords:** Viktor Orbán, political communication, state of the nation address, automated text analysis, fear, enemy image

## 1 Introduction

The media frequently use fear as a tool because of its attention-engaging nature (Altheide, 2013). It is also commonly used to steer the focus of a population away from certain topics (Sik, 2020). Hungarian politics has, in recent years, become filled with fear due to ever more regularly emerging enemy images created by the government. The right-wing ruling parties of the country have built up a nationalist, exclusivist system that is strongly based on narratives of war and the need for the protection of the state from ‘others’ (Demeter, 2018).

---

According to Gerő et al. (2017), the negative marking of vulnerable social groups in Central and Eastern Europe is now more present than at any other time since the Second World War. Populist parties in Poland and Hungary have gained substantial popularity by using images of both insider and outsider enemies and stressing the importance of national sovereignty (Csehi & Zgut, 2020). While the regimes of the region differ in certain aspects, populism certainly plays an important role and has commonalities within Central-Eastern Europe (Kopper et al., 2023). Populist fearmongering is, therefore, a crucial issue of significant relevance to understanding contemporary Hungarian and Central-Eastern European politics.

The trend to large-scale fearmongering is in line with the Schmittian friend-enemy dichotomy, which suggests the necessity of the existence of groups of 'strangers' to enable identity formation against them. These enemies strengthen the 'insider' group and represent harmful elements that threaten society. As weak liberal societies are unable to protect themselves against their enemies, groups that clearly point out who the enemies are, have a substantial advantage, according to Schmitt's theory. These strong groups can be seen as potent candidates for tackling the threat of the stranger (Schmitt, 1932/2008). Hungarian and Polish populists have successfully portrayed the European Union as a weak system that is unable to overcome its challenges (Kopper et al., 2023). This weak liberalism is tackled on a Schmittian basis by the 'illiberalism' of Viktor Orbán, which is posed as a strong, reliable system (group) that knows its enemies and has its ways to fight them.

The Hungarian Prime Minister, Viktor Orbán, is a key figure in the country's public life and has an important role in European politics as well. His populist communication strategy has led to significant changes in Hungary and attracted numerous international followers (Körösényi & Patkós, 2017). With his system of 'illiberal democracy', he has been a model for many right-wing parties in the region (Krekó & Enyedi, 2018). He has supported candidates in the national elections of Serbia, Slovakia, and Slovenia, creating international bonds and forming alliances with like-minded politicians. The State of the Nation addresses of Viktor Orbán are considered cornerstone events in which he reports on the achievements of the past year and outlines his plans for the year ahead. The political content of these speeches is, therefore, of paramount importance, which is why we focus on these texts in our research.

The images of the 'enemy' have played a significant role in the political discourse of the last decade in Hungary. Whether in the media, on billboards, or in National Consultations, they have become part of public narratives and strong reference points in political communication (Kitzinger, 2000). By examining the different enemy images originally defined in the National Consultations and Viktor Orbán's speeches, we can map the origin of enemy images and their propagation patterns.

The study is intended to show how official communication is used for fearmongering in Hungary through the example of the prime minister's most prominent speeches. The main research question was as follows: How are the enemy images identified in the 2015-2022 National Consultations echoed in Prime Minister Viktor Orbán's State of the Nation addresses? Our empirical analysis relied on both quantitative and qualitative methods; we used critical discourse analysis and automated text analytics to study the enemy images in the prime ministerial speeches. We computed the weighted relative frequencies of target words associated with depicting enemy images and applied statistical tests to check whether they were significantly more or less frequent in different time periods and years.

We also examined the keynote speeches delivered on the anniversaries of national holidays and compared the patterns observed in them to those revealed in State of the Nation addresses.

Our results indicate that there were five different enemy images used for attuning listeners to fear ('Brussels', 'migrants', 'George Soros', 'Ferenc Gyurcsány', and 'opposition') in the State of the Nation addresses, which were frequently referred to in other types of prime ministerial speeches as well. The temporal dynamics of the emergence of these enemy images are in line with our prior assumptions, supporting the insights from our qualitative analysis.

## 2 Theoretical background

### 2.1 Fear and moral panic

Furedi (2006) argues that fear plays a central role in our societies. The culture of fear involves an obsession with safety in a society that has not experienced truly fearful events, such as wars. Fear in current Western societies may often seem baseless regarding the everyday reality. However, it is present and expands, especially through the media, without firsthand experience of real terrifying events. Fear is not as much an individual emotion as it is a constant state of being in the world. It is a form of vigilance that is largely based on the moderated content that one receives and accepts as a representation of reality (Furedi, 2006).

According to Domonkos Sik (2019), this culture of fear derives from a crisis of integration. In such a culture, individuals cooperate based on a fear of becoming 'the other'. However, these latter groupings, which are not based on honest solidarity, are much more prone to identify outsider enemies in relation to whom the group can define itself and strengthen cohesion.

Stanley Cohen's theory of 'moral panic' (1972) fits well with the culture of fear. Moral panic implies the fear of the destruction of society's shared culture by an outsider enemy. The creation of moral panics is not a shallow process. They are efficient because they are built on deeply held stereotypes or assumptions and involve a convincing theory about why bad events played out the way they did. This involves enemies that are fearful and threaten the very foundations of social order. The tactic of deploying moral panic in a society that is obsessed with safety can generate great political advantages.

Viktor Orbán's campaign strategists found moral panic to be a useful tool for strengthening the government's position. This is the basis of what Endre Sik (2016) calls the 'moral panic button' – a campaign strategy that involves the creation of moral panics and uses the fear thereby generated to strengthen the in-group, thus maintaining the government's popularity. The strategy resembles a button, the pushing of which results in an immediate moral-panic-generating campaign that reaches every corner of the country. The usage of the moral panic button creates increasing in-group/out-group division, which needs to be maintained if the government is to keep its support – however, it can lead to the enhanced control of public discourse and ultimately to the decline of the institutions of democracy (Gerő & E. Sik, 2020).

The culture of fear is universal in contemporary Western societies (Furedi, 2006); however, the systematic utilization of this for fearmongering through moral panics is exceptionally severe in Hungary (Sik, 2016). Meanwhile, enemy creation has also been used as a tool in other Central-Eastern European countries (Spasojević, 2019). Supplemented with the strong presence of right-wing populism in the region (Kende & Krekó, 2020), pressing the moral panic button may become a prominent choice if other governments are willing to follow the Hungarian example. This could lead to more division and societal disintegration in Central-Eastern Europe.

## 2.2 Enemy images

Holding opposing views and having political opponents is a crucial aspect of democracy; an opponent, however, is not equivalent to an ‘enemy’ (Szabó, 2007). The tendency to turn opposing parties into enemies has increased in frequency in a few post-socialist countries that are drifting away from liberal democracy (Szelényi & Csillag, 2015). Hungary has, in recent years, become one of the most notable examples of such states. Since Viktor Orbán’s election win in 2010, the country has taken a turn towards an ‘illiberal’ system (Szelényi & Csillag, 2015), which, among other elements, is constructed through the constant maintenance of enemy images (Kopper et al., 2017). Fear is a basic emotion that is a useful instrument for defining in- and out-groups, as well as shifting the focus away from certain issues (D. Sik, 2020). It has long been used in Hungarian politics to divide people. While the left created a discourse in which the right was pictured as the ‘enemy of democracy’, in the rightist discourse, the opponents were the ‘enemy of the nation’ (Kopper et al., 2017). These divisive narratives have accelerated in recent years, especially since the refugee crisis of 2015, which stirred up xenophobic attitudes throughout Central and Eastern Europe (Walter, 2019). According to Kopper et al. (2017), Viktor Orbán has been very successful in creating fearsome enemy images, in opposition to which he can place himself in the position of a ‘protector’ with the ability to save the nation from the outside and inside threats.

The presence of enemy images in Hungarian politics can be observed not only in political speeches but also on billboards, media advertisements, and in the news. Orbán’s government has successfully built a hegemony (Cooper, 2023) that involves control over a large part of the media sector, enabling it to spread and form messages that support its standpoint (Bajomi-Lázár, 2013).

These enemy images often serve to promote group cohesion and distinguish ‘us’ from ‘them’. This has resulted in Viktor Orbán being able to unite his community in a way that other politicians have been unable to in recent years (Gerő & Szabó, 2017). Hence, using enemy images goes beyond gaining popularity by pointing to one stranger or another. It advances the feeling of belonging to the in-group, as the out-group is clearly defined.

## 2.3 National Consultations

Public consultations about policy changes or important initiatives can be identified in several countries throughout Europe (Pócza & Oross, 2022). They take various forms. Some are organized bottom-up, initiated by citizens or non-governmental organizations, while

others are arranged top-down by governments (Gáspár, 2021). According to Pócza and Oross (2022), such popular consultations can be useful components of a deliberative democratic system, as they allow citizens to participate in nationwide or local decision-making processes. The Hungarian example of a popular consultation is, however, different from other similar initiatives in Western Europe. The Hungarian so-called 'National Consultation' is much more of a marketing tool than a representative opinion poll (Gerő & E. Sik, 2020). Fidesz's top-down, government-organized consultation politics started in 2010 when they won the general election. As Endre Sik (2016) puts it, National Consultations are used periodically to create moral panic and effectively manipulate voters. The questions in Hungarian National Consultations are often leading; phrased in a manner that does not promote objective reflection, and the scientific professionalism of surveying.

Enemy images are crucial elements of National Consultations. In his theory about the moral panic button, Endre Sik (2016) describes a campaign process used by Orbán's government that is largely built on the National Consultation. He argues that when the government is in need of popular support, it creates an 'enemy' that is depicted as dangerous to the nation and launches a vast campaign against it. The script of these campaigns includes easy-to-remember slogans, media, and street advertisements, which are disseminated as official information. Such 'information campaigns' are then concluded with a National Consultation, followed by another campaign informing the public about the consultation's results. The ultimate goal, as Endre Sik (2016) sees it, is the creation of moral panic, which is a fear of a danger that may allegedly destroy the given society's shared cultural and moral heritage (Cohen, 1972). National Consultations are, therefore, efficient and frequently used by the Hungarian government to create fear and manipulate citizens. Over the past few years, they have proven to be a sufficient means of building popular support before elections or other times when needed (Gerő & E. Sik, 2020).

This research addresses the National Consultations in the period between 2015 and 2022. Although there were National Consultations before 2015 as well, the choice to start the examination in 2015 was based on the theory of the moral panic button, which was systematically implemented in the government's strategy in 2015 to restore Fidesz's declining popularity (Gerő & E. Sik, 2020). Within this time frame, eight consultations featured numerous 'enemies'; however, examining the surveys made it clear that a few of these enemy images have been present in multiple National Consultations. These reoccurring enemy images became the focus of this research.

The most frequent enemies of the National Consultations have been (1) 'Brussels', which in Orbán's narrative refers to the European Union, (2) immigrants, and (3) George Soros. Orbán has long fostered mistrust of the European Union and has several ongoing disputes with its leaders (Scheppele, 2022). In government advertisements, campaign materials, and speeches, if the European Union is mentioned in a negative context, it is denoted by the name of the Belgian capital, Brussels. The topic of migration, the second most frequently deployed enemy, is crucial for Viktor Orbán. According to Scheppele (2022), this has brought him numerous voters and, from time to time, helps to ensure his election majorities. The American-Hungarian billionaire George Soros, the third most important enemy, has become a central enemy image, especially because of his support for the integration of immigrants (Reményi et al., 2023), but his vision of Open Society is also the absolute opposite of Orbán's illiberal approach (Kopper et al., 2017). In the right-oriented media, Soros has been connected to support for illegal migration and civil associations that criticize the



Hungarian government. These three most repeated enemy images (Brussels, immigrants, and Soros) are accompanied by several other 'shorter-lived enemies.' Further, there are at least two more enemies whose presence has not been particularly striking in the National Consultations, but based on their appearance in the media, advertisements, posters, and speeches, are worth mentioning and analyzing. These are (4) Ferenc Gyurcsány, the former Prime Minister, whom Orbán likes to present as his archenemy, and (5) the 'opposition' (most opposition parties, regardless of political orientation). In the current political discourse, these five enemy images seem to be most frequently represented based on the National Consultation, news, advertisements, and media presence. Therefore, these are at the core of this study.

## 2.4 State of the Nation addresses

Every February, Orbán gives a speech in which he reports on the previous year's achievements and plans for the coming year. In addition to members of the press, these speeches are typically attended by high-ranking public figures, such as the president and members of the government. The primary target audience is, therefore, typically highly educated, pro-government intellectuals. The annual State of the Nation speeches are broadcast live to the public, usually on HírTV, a channel close to the government. In this study, we focus on the State of the Nation speeches between 2010-2022, which covers the second period of Orbán's prime ministership until the end of the study. Except for the first speech in 2010, all of them were delivered by Viktor Orbán in his capacity as Prime Minister. The key element of these speeches is a summary of the state of the country. In the beginning, this involved a stronger economic focus; however, even then, there were a few speeches intended to demonstrate strength. 'Fighting as a metaphor [...] has been reintroduced into political rhetoric and thus into political discourse (e.g., anti-EU "freedom fight", "utility-price fight", "We will not be a colony" etc.)' (Tóth, 2015, p. 117). Such militant, revolutionary statements and speech styles have been associated with Viktor Orbán throughout his political career. His speaking style has, in the meanwhile, undergone several changes. As Tóth (2015) observes, at the beginning of his career, Orbán stood out among the politicians of his time for his complex yet easy-to-follow sentences. Later, however, this changed, according to Tóth (2015), and less complex, shorter messages became more typical. Another change is an increasing emphasis on nationalism in Orbán's communication. The nation's unity beyond borders became part of the speeches explicitly and in jokes and proverbs, too (Tóth, 2015).

During the qualitative analysis of the speeches, other changes were also noticed. The tone, which we initially categorized as 'civic-conservative', gradually changed, accelerating pace after 2014 to a more populist, more emotional tone. The related speeches, along with the simplification in style and increase in emotion-provoking content, more often included information from unclear sources, often categorized in public discourse as conspiracy theories. A good example of this new tone is a passage from the 2022 State of the Nation address: 'If we let the pro-migrant Brusselian bureaucrats help the both ridiculous and dangerous characters of the Gyurcsány-show to win, they will open the borders' (Viktor Orbán's State of the Nation address, 02/12/2022). This quotation is worth highlighting for several reasons. First, it presents three enemy images (immigrants, Brussels, and

Gyurcsány) in a single sentence. Second, in an unprecedented way, it refers to the exact components of the Fidesz campaign that were shown in the media and on posters. Third, it presents an unverifiable piece of information as a clear truth. In addition, attention should be paid to the stylistics and the simple, consistent wording of the sentence, which makes an equivalence between the enemy images and an undesirable event presented as a threat.

The interest of this research, apart from Viktor Orbán's State of the Nation addresses, was other keynote and annual speeches of his. These were chosen based on the attention paid to them by the media and the size of the audience that was reached. We analyzed the prime ministerial speeches from the national days of March 15 and October 23, as well as the regular lectures given by him at the Bálványos Free Summer University and Student Camp in Tusnádfürdő, Romania.

### 3 Data and research questions

The examined Hungarian text corpus was prepared for the research project of ELTE Research Center for Computational Social Science<sup>1</sup> (RC2S2), entitled 'The layers of political public sphere in Hungary (2001–2020)'.<sup>2</sup> The corpus contains texts collected from the miniszterelnok.hu website and its archive. The raw texts in the corpus were collected and pre-processed by Zsófia Rakovics, a researcher of ELTE RC2S2.

The database of speeches also includes metadata such as information about the circumstances in which the speech was given, for example, publicly available information about the date and place of the speech, the official title of the speech, and a short introduction (lead) to the speech. Given that the dates of the speeches were available, we created a year variable and used this for aggregation during the analysis. We also visualized the relative frequency of relevant keywords related to enemy images according to time (by using the year variable) and created an index to signal the time of the parliamentary elections, which was also used for the visualization. We categorized the speeches by their type and identified State of the Nation addresses (the main target of our analysis) and other keynote speeches, which included the prime ministerial speeches of the national holidays of March 15 and October 23 and Viktor Orbán's annual lecture at the Bálványos Free Summer University. The keynote speeches were chosen based on their recurring nature and the media's strong attention to them.

Our research sought primarily to identify the way in which the enemy images in the National Consultations appear in the State of the Nation address speeches of Prime Minister Viktor Orbán. Furthermore, we aimed to investigate the presence of these enemies in other keynote speeches and compare the frequency of their mention to those in the State of the Nation addresses. As a result, the following research questions were formulated for the study:

---

<sup>1</sup> The website of ELTE Research Center for Computational Social Science (RC2S2): rc2s2.elte.hu

<sup>2</sup> Details of the research are available on the RC2S2 research group website, under the research projects section 'The layers of political public sphere in Hungary (2001–2020)'. The research was funded by NKFIH, with the following research ID: K-134428. The principal investigator of the project is Renáta Németh.

1. How are the enemy images identified in the 2015-2022 National Consultations echoed in Prime Minister Viktor Orbán's State of the Nation addresses?
2. How do these same enemy images appear in other keynote speeches?
3. Is there a difference in the frequencies of appearance in the two different types of speeches?

Aligned with the main research questions, we filtered for the relevant types of speeches, namely State of the Nation addresses, and focused on those texts, specifically examining the full corpus (all available transcripts) to draft an overview of the talks of Prime Minister Viktor Orbán.

The hypothesis was that enemy images would appear in both types of speeches (the State of Nation addresses and the keynote speeches) but be associated with different patterns as the audience also differed. Based on the moral panic button theory of Endre Sik (2016), it was assumed that the dominant periods of different enemy depictions would occur at different times. On the basis of the National Consultations, it was expected that the mention of 'immigrants' would peak in 2015, 'Brussels' in 2016, 'George Soros' in 2017, while 'Ferenc Gyurcsány' around 2010 and the 'opposition' around 2022.

## 4 Methods

For the research, we applied qualitative and quantitative methods, using critical discourse analysis and automated text analytics to study the enemy images in prime ministerial speeches. The qualitative analysis enabled the examination of implicit fearmongering, while the quantitative analysis provided data regarding the explicit appearance of enemy images. The utilization of the two methods resulted in a more holistic understanding of the enemy images that were deployed.

We defined targeted keywords according to the focus of our research and combined words with the same substantive content (in other words, we searched for synonyms). We categorized and interpreted these words in the 2010-2022 State of the Nation addresses by annotating them to examine their meaning qualitatively. As a result, the database built for the analysis included not only the metadata of speeches and the raw texts but also the occurrences of individual words within each text (so-called document term frequency). Utilizing basic methods of automated text analysis, we used bag-of-words models and examined the absolute and relative occurrences of keywords in the speeches. To compare the results statistically according to years and type of speeches, we applied statistical tests. The findings are documented in the *Results* chapter.

### 4.1 Qualitative analysis

We studied the National Consultations qualitatively and then created a database of words categorized as enemy images, in which the enemy images were the variables sorted by column, while those in the National Consultations were sorted by row. In this database, a '1' denoted if the enemy image appeared in the current consultation sheet and a '0' if it did not. As discussed above briefly, the dataset and the concept of the analysis were later simplified by combining words with the same meaning or referring to the same enemy.

**Table 1** Representations of enemy images in the National Consultations.  
The total number of National Consultations examined is equal to 8 (N=8)

Enemies	Frequency	Frequency in percentage (%)
Brussels	7	88%
Immigrants	5	63%
George Soros	3	38%
Socialist-liberal party coalition	1	13%
Amnesty International	1	13%
Helsinki Committee	1	13%
NGOs supported internationally	1	13%
ISIS	1	13%
Terrorism	1	13%
Political parties of the opposition	1	13%
Banks	1	13%
Ferenc Gyurcsány	1	13%

The table above shows the results of the analysis of the dataset. We can find in each row the enemies named in the National Consultations and the corresponding frequency. The list of enemies includes not only people (George Soros, Ferenc Gyurcsány) but also organizations (Amnesty International, Helsinki Committee, NGOs, banks, a party coalition), the name of a city (Brussels, referring to the European Union), and concepts (immigrants, terrorism).

Eight National Consultations took place in the period under review, of which Brussels appeared as an enemy in seven. The only omission was the online-only survey 'About a New Start' in 2021. The second most frequently mentioned enemy was immigrants, a group that appeared in five questionnaires in total. The third most frequently mentioned enemy was George Soros, whose name was mentioned in three National Consultations. The remainder of the enemies appeared in only one questionnaire each, but this does not mean that they are short-lived, one-time-only enemy images, as the Prime Minister's speeches attest. Ferenc Gyurcsány, for example, was only named in one questionnaire, yet he is recurrently depicted as an enemy in the rhetoric of Viktor Orbán and Fidesz-KDNP. For this reason, we compared the underrepresented actors in the National Consultation with their appearance in the campaigns. We found the five most typical enemy images to be Brussels, George Soros, Ferenc Gyurcsány, the opposition, and migrants. Gyurcsány, a former Prime Minister, is a dividing personality and is widely unpopular among many Hungarians, regardless of their political standpoint (Gerő & Szabó, 2017). Connecting his person to the whole opposition and pointing out this important enemy was a campaign strategy that promised success; therefore, he was not left out of this analysis either. The political opposition is the natural opponent of the government, which has also been turned into an enemy in the Hungarian case. Based on this, it was also considered important to

include the opposition in the analysis beyond its frequent representation in campaign materials. Therefore, the study analyzed the occurrence of the five abovementioned enemy images.

In the qualitative research, we started by using Fairclough's principles of critical discourse analysis (1995; 2013). An important element of this is the recognition of the ideological use of language (Keller, 2012), so we took these guidelines into account throughout the research process and in interpreting the results.

The keynote speeches were chosen based on their recurring nature and the media's strong attention to them. In this research, we declared the prime ministerial speeches given on the national holidays of March 15 and October 23, along with Orbán's annual lecture at the Bálványos Free Summer University, to be keynote speeches. Hungarian media tends to closely follow these events as they can often be regarded as crucial, course-setting talks, shedding light on the future direction of the government. The State of the Nation addresses and the abovementioned keynote speeches are not only of high importance but are repeated annually in similar settings. Hence, they provide an outstanding opportunity for analysis and over-time comparison. The observation of enemy depictions in these speeches helps illustrate the overall relevance of enemies throughout time in the narrative of the current Hungarian Prime Minister.

We examined the current Hungarian Prime Minister's State of the Nation address speeches more thoroughly by annotating them qualitatively. In this process, we created categories based on speaking style, substantive content, and the main message of the speeches. The goal was to examine possible changes in style over time and to observe how these official speeches were influenced by the campaign mechanisms.

The discourse analysis approach enabled the examination of underlying, quantitatively undetectable messages present in the speeches. The linguistic nuances and sentence formulations could be highlighted, making it possible to detect implicit fearmongering and changes in tone of voice.

## 4.2 Quantitative analysis

In the quantitative part of our research, we examined the appearance of the previously defined keywords representing different aspects of the various enemies mentioned in the State of the Nation addresses. We also studied the keynote speeches given on the anniversaries of national holidays and compared the patterns we observed to those revealed in the State of the Nation addresses.

Before starting the quantitative investigation, we calculated the length of the texts, which is defined as the sum of the occurrences of the words that make up the speeches. We used this information to compute relative frequencies from the observed frequencies by dividing the observed frequency of the target words by the length of the actual speech and then multiplying by one hundred to filter out variations in the number of words in the texts before the analysis. Thus, we computed the weighted relative frequencies of selected keywords associated with the depiction of enemy images and applied statistical tests to check whether they occurred significantly more or less frequently in different time periods and years.

We followed the methodology used by Boda and Rakovics (2022) when analyzing the relative frequency of migration-related keywords in prime ministerial speeches. We examined the average occurrences of the most typical words associated with enemy images in the highlighted speeches (both State of the Nation addresses and different types of keynote speeches) using independent sample t-tests and then visualized their distributions by including the time dimension in the analysis. We also used paired-sample t-tests to compare the relative frequencies of the different keywords that occurred in the speeches. Further, most importantly, we examined the results over time to generate information about the dynamics of change over the years. The latter are detailed in the *Results* chapter of our article.

We hypothesized that based on the peculiarities of the State of the Nation address speeches – namely, those whose audience is a rather limited, politically highly engaged, well-educated group – they would feature enemy images less frequently than the keynote speeches that are directed towards the wider population. Therefore, we expected to identify fewer words connected to inimical characters in the State of the Nation addresses and more in the other keynote speeches.

## 5 Results

### 5.1 Qualitative results

By applying qualitative methods, we learned about the meaning and interpretation of the examined keywords and the substantive context in which they occur. During the annotated interpretation, we noticed that the enemy images from the National Consultations appeared in higher proportions than previously expected, according to our hypotheses.

The narrative of the 2010 State of the Nation address was dominated by an emphasis on the need for strong government. It is important to note that this speech was delivered during the government of Gordon Bajnai, so Viktor Orbán was still speaking as a politician in opposition and as the prime ministerial candidate of Fidesz-KDNP. In this State of the Nation address, he essentially summarized the policy guidelines for the coming years, which he would steer. He set out the main principles of a 'work-based society', talked about the policy of opening to the East, and criticized the work of the government of the time, which he described as weak. His main message was to 'rebuild' the country and put it on a strong national foundation.

Putting the country on a new foundation was also a central theme of the 2011 State of the Nation address speech. Here, in addition to outlining the economic situation and political goals, images of the enemy were more prominent. The main enemies of the period were 'speculators', but these speeches never declared exactly who they were, and no clear information was shared about them, only that they posed a threat to Hungary. However, the radical communication offensive against these mysterious groups that we learned about later was not yet launched in the current Prime Minister's speeches.

In his 2012 annual assessment, Viktor Orbán declared that 'Hungary stands on new foundations'. This speech also had a stronger economic focus and discussed plans, a new constitution, and workplace creation. There was also criticism of the European Union,



with the claim that 'Europe is slowly becoming like alcohol: it inspires great goals and prevents us from achieving them' (Viktor Orbán's State of the Nation address speech, 02/07/2012).

The State of the Nation address of 2013 was the first time that recurring slogans appeared that we had seen in various government campaigns. The key message of this year was 'Hungary is doing better' (In Hungarian: 'Magyarország jobban teljesít'). This speech was already intended somewhat in preparation for the 2014 elections, and the word 'socialists' as a source of danger was mentioned several times.

In 2014, an increase in references to religion was noticed. Viktor Orbán quoted from the Bible twice and used religious terms several times. In addition, the later to be emblematic 'Brussels bureaucrats' appeared here, and there was a general shift in terminology. From this point onwards, the European Union was increasingly consistently referred to as 'Brussels'. Left-wing political parties and socialists are also mentioned in a negative context, which is not surprising given that the speech was made in the run-up to the parliamentary election campaign.

The year 2015 was an important turning point in Viktor Orbán's communication. As Gerő and E. Sik (2020) wrote, the most unfavorable opinion poll for Fidesz-KDNP came out in January of that year. There was a drastic drop in the party's popularity, which was responded to with a radical communication strategy. Institutionalized fearmongering was, from then on, part of campaigning (Gerő & E. Sik, 2020). This part of the 2015 State of the Nation address is an example:

'Migrants from the west of the continent are being recruited by terrorist organizations, while the southern borders of the EU, including our own, are besieged by a new era and wave of migration, which is baffling increasingly frustrated states and governments. And all this is happening at a time when millions of Western Europeans feel they have to work more and more for less and less money if they are able to keep their jobs at all.' (Viktor Orbán's State of the Nation address speech, 02/27/2015)

The speech showed growing populist overtones, and for the first time, climate change was mentioned as a problem to be tackled. Endre Sik (2016) also dates to this year the construction of the previously mentioned 'moral panic button', which has defined the political discourse in Hungary until the present time and successfully restored the loss of popularity of Fidesz-KDNP.

In the 2016 State of the Nation address, we hear a more populist Viktor Orbán who is more likely to base his message on emotions and fears. 'We will teach Brussels, the people smugglers and the migrants that Hungary is a sovereign country and that the only way to enter its territory is to obey our laws and obey our uniformed officers' (Viktor Orbán's State of the Nation address speech, 02/28/2016). The new slogan was 'stopping Brussels', and he assured his audience that 'Hungarian reforms are working' (Viktor Orbán's State of the Nation address speech, 02/28/2016).

In 2017, the emergence of enemy images became even more pronounced than before: migrants, George Soros, and the decline of the European Union were discussed. It is striking that in this speech, there was less talk of the economy, which was the focus at the beginning, and of the reforms that had been carried out and planned. The dominant narrative typically involved fighting enemies. It is also interesting that, in this speech, he used whole sentences from his annual assessment from the previous year.

In the heat of the election campaign, the 2018 State of the Nation address was not short of enemies. Viktor Orbán, who described himself as the defender of Christian Europe, talked about the network of Soros, the Soros plan, immigrants, and an opposition unfit to govern. 'Stop Soros' appeared as an emblematic slogan, and we can observe the growing emphasis on us-versus-them groupings (Molnár, 2020).

In the 2019 State of the Nation address, there was a renewed emphasis on economic growth, but enemies were not pushed to the background either. The Prime Minister analyzed in detail the incompetence of the political parties in opposition and ridiculed them, for example, by declaring, 'all we can say about our fellow Members of Parliament in the opposition is that they are indeed every inch people of courtesy and refinement: every inch up to their ankles, but no further!' (Viktor Orbán's State of the Nation address, 02/10/2019). Brussels was not left out either since it was claimed that 'in Brussels, the seven-point work plan has already been drawn up and is ready for deployment, with which they want to turn the whole of Europe into an immigrant continent after the European elections' (Viktor Orbán's State of the Nation address, 02/10/2019). It is important to note that by 2019, several National Consultations and billboard campaigns had already been implemented, in which the images of the enemies that appear here were central figures. In this speech, the person of George Soros was presented as a mystical background manipulator of power who strongly wanted Hungary's downfall.

In Orbán's 2020 State of the Nation address, he talked about Soros, Brussels, and the opposition was also mentioned in the populist, fighting-for-freedom tone familiar from recent years. There was also much talk of the Trianon anniversary and the tenth anniversary of the Fidesz government. Orbán said: 'The facts [...] show that the last ten years have been the most successful ten years in the history of Hungary over the last hundred years' (Viktor Orbán's State of the Nation address, 02/16/2020). He also noted that he believed that democracy based on National Consultation was 'true' democracy.

As the State of the Nation Address of 2021 was canceled due to the COVID-19 pandemic, the traditional February annual assessment speech was not held again until two years later, in 2022. Given that this was again in the parliamentary election campaign period, it was not surprising that there were references to the political parties in opposition, this time mostly mentioned as 'the left'. That was the last speech that was analyzed in our research and may be considered a summary of all previous addresses of the State of the Nation. It was the most radical of all. All the enemies of previous years appeared in it, as well as the terrifying shadow of war. Orbán also declared that 'the stakes of the election are huge; it is a vital turning point for Hungarians' (Viktor Orbán's State of the Nation address speech, 02/12/2022). This speech was closer to being a campaign speech than an annual review in the sense of 2010. Images of enemies and fearmongering played a key role in the narrative, which was concluded with: 'Saddle up, the campaign has begun, it's time for us to ride out. We'll be back in fifty days, from the right hand's side! God above us all, Hungary first! Go, Hungary, go, Hungarians!' (Viktor Orbán's State of the Nation address speech, 02/12/2022).

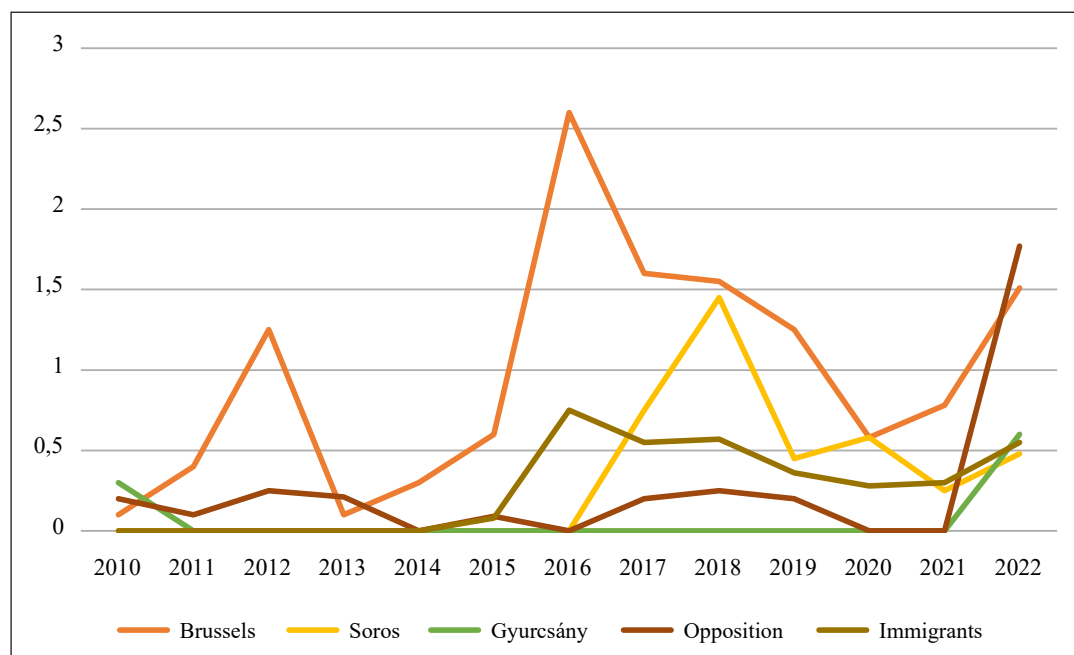
In conclusion, the qualitative analysis helps identify a radicalization of the way of speaking in the State of the Nation addresses. Images of the enemy always appeared in the speeches but were particularly emphasized after 2014. The initial function of the State of the Nation addresses gradually changed from 2016 onwards, at an accelerated pace, into a

campaign of fearmongering and maintenance of the fear induced by political communication. Thus, the strategic, ideological use of the speeches (Fairclough, 2013) increased compared to the context-following functional use that was typical before.

## 5.2 Quantitative results

During the quantitative analysis, we studied how the relative frequencies of the various images of the enemies identified in the National Consultations changed over time. We also studied the State of the Nation addresses separately, and together with the keynote speeches of the current Hungarian Prime Minister, the talks of March 15, October 23 and at Tusnádfürdő.

Focusing on the changes over time, we created graphs of the temporal variation in the appearance of the five identified enemy images, as detailed earlier.

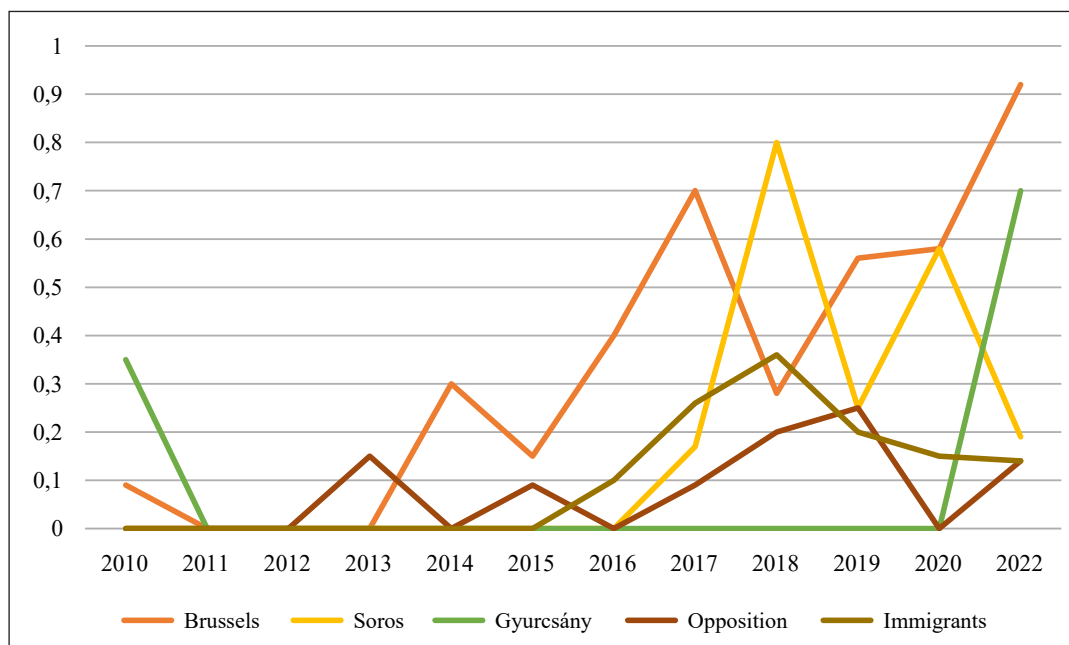


**Figure 1** Relative frequencies of mention of the five enemies over time. The horizontal axis shows the years, and the vertical axis shows the relative weighted frequencies. The graph shows the relative frequencies in all types of prime ministerial speeches according to time for the words 'Brussels', 'Soros', 'Gyurcsány', 'opposition', 'migration'.

Figure 1. shows the occurrence of words related to images of the enemy in all the featured speeches by year. We used standardized variables for measuring relative frequencies, created as described in the methodology section. There was relatively large variation in the mention of enemy images. 'Brussels' appears to be the most frequently used in prime

ministerial speeches. Gyurcsány, the political parties in opposition, and Brussels were mentioned in 2010 as enemies in talks. The relative weighted frequencies of Brussels increased and reached a local peak in 2012. In 2013, the relative frequency of the occurrences of the 'opposition' was greater than that of the other enemy images, but from 2014 onwards, Brussels became the most often discussed enemy of Hungary, according to Viktor Orbán. The weighted measures of Brussels, referring to the European Union, increased between 2013 and 2016. The latter year – 2016 – is important because a new enemy appeared: migration, which occurred relatively frequently. From then onwards, mention of migration and Brussels became relatively frequent in the speeches. Mentions of migration reached a local maximum in 2016, and after a slight decrease (in 2017), saw an uptick in 2018 as well. In the same year (2018), the enemy image of George Soros was used by Viktor Orbán relatively frequently, almost as often as Brussels. In 2020, the relative dominance of both Brussels and Soros was equal in the prime ministerial speeches. In the year of the parliamentary elections, the relative frequency of mentions of all enemies increased, which finding aligns well with the results of the qualitative analysis and the fact that 2022 was a year of parliamentary elections. Mentions of the opposition and Ferenc Gyurcsány became dominant in that period, while Brussels remained the number one enemy according to the State of the Nation addresses and the keynote speeches of Viktor Orbán.

In the second graph, Figure 2, we have focused closely on the State of the Nation addresses and plotted the emerging enemy images by year, similar to Figure 1.



**Figure 2** Relative frequency of mentions of the five enemies over time in the State of the Nation addresses. The horizontal axis shows the years, and the vertical axis shows the relative weighted frequencies. The graph shows the relative frequencies in the State of the Nation address speeches according to time for the words 'Brussels', 'Soros', 'Gyurcsány', 'opposition', 'migration'.

Please note that, as no State of the Nation address was held in 2021, we have no data available for that year.

The results of Figure 2 are slightly different from those of Figure 1: a shift in time is observable in the appearance of enemy images and a decrease considering the volumes – the relative frequencies – over time. The dominance of the different enemies changes more drastically in the State of the Nation addresses when considering time. This finding aligns well with the expectation that the role of the State of the Nation addresses and the other keynote speeches (March 15, October 23, and Tusnádfürdő) are conceptually different.

Brussels is again one of Hungary's most dominant enemies, but the relative frequency of mentions of this is not that high throughout the examined period. The peaks of relative frequencies for Brussels are in 2014, 2017, 2019 and 2022. Taking the date of the parliamentary elections into consideration, it is quite visible that the word 'Brussels' was used at the most critical time periods when the elections took place. The dominance of 'George Soros' is greater in the State of the Nation addresses than observed in all targeted speeches (including the keynotes as well). Viktor Orbán depicted Soros as an enemy most dominantly in 2018 and 2020. Ferenc Gyurcsány, the former Prime Minister of Hungary, was mentioned most frequently in 2022, in the year of the parliamentary elections.

According to the State of the Nation addresses, the enemy images in each year of the elections were dominant to varying degrees. In 2010, Ferenc Gyurcsány was the most frequently mentioned enemy, replaced by Brussels in 2014, the only enemy in that year. George Soros became the number one enemy in 2018, with the highest relative frequencies in that year, when migration, Brussels, and the political opposition were also mentioned relatively frequently in the State of the Nation addresses. In 2022, Brussels and Ferenc Gyurcsány became the most often mentioned enemies of the country.

The observations based on the quantitative methods that were used therefore confirmed the qualitative results. In other words, the images of the enemy that appeared from the outset were radicalized after 2014 and became an increasingly integral part of Viktor Orbán's narrative. Finally, 2022 was the year when most enemies appeared at the same time, as seen in the figures (Figures 1 and 2). This is the only year in which all the enemy images examined are represented in the State of the Nation address speeches.

## 6 Discussion

Based on the reviewed literature, it was apparent that an increase in the presence of enemy images in the speeches of the Hungarian Prime Minister was to be expected. This was confirmed by both the qualitative and the quantitative analysis.

In the paper, we have pointed out that enemy images were already part of Viktor Orbán's narrative in 2010, in the first examined speeches, but only in 2014 and around 2015, when the National Consultations began to become more serious, did they explicitly start to gain ground. Enemy images of individuals also appeared at different times. Brussels and Ferenc Gyurcsány were the earliest enemy images in the State of the Nation addresses, while all the keynote speeches also mentioned the opposition. This latter group only appeared in 2013 in an annual State of the Nation speech. Mention of immigrants started in 2015 in the keynote speeches and became part of the narrative of the State of

the Nation addresses in 2016, while George Soros appeared in both types of speeches in 2017. Enemies were also a dominant feature of the State of the Nation addresses in election years. Specifically, in 2018, several inimical figures were mentioned, as well as in 2022. The last annual review speech that we examined, from February 2022, proved to be the most radical. This was confirmed by both the results of the qualitative interpretation and the quantitative statistical analysis. These revealed all the identified enemy images, accompanied by a new threat – war. This speech can be considered part of the 2022 election campaign, as it contains several elements related to it. Based on Fairclough's (2013) model of critical discourse analysis, the speech can be assumed to transmit ideological content.

The findings reinforced the pattern that was established by previous studies in the field (Demeter, 2018; Gerő et al., 2017; Gerő & E. Sik, 2020; Kopper et al., 2017; E. Sik, 2016), namely that enemy images are beyond doubt present in Viktor Orbán's speeches, and a radicalizing tendency can be observed alongside a seemingly constant change in the nature of the enemies. This research thus adds value to the examination of discourse with its use of quantitative tools that enabled the statistically supportable examination of changes in enemy images. The simultaneous use of discourse analysis generated a deeper understanding of these changes, reflecting on the implicit, underlying enemy creation. Similar methods have earlier been applied by other studies in the analysis of Orbán's communication on migration (Boda & Rakovics, 2022; Sylá, 2023); however, utilizing these methodologies to identify enemy images in important official speeches over a longer period of time is the novelty of this study.

We observed that the presence of certain enemies in the Prime Minister's speeches correlates with the patterns of enemies specified in Márton Gerő and Endre Sik's (2020) study about the pressing of the moral panic button. However, the mention of these enemies might peak in frequency at slightly different times in the speeches. Meanwhile, it is clear that the panic-generating images of the National Consultations found their way into the keynote and State of the Nation speeches.

Many scholars have conducted research on Viktor Orbán's speeches before us. Comparing these previous pieces of work to our own, it is apparent that the enemy images have changed over time. This suggests that the frequency of the observed enemies differed over time, as did the enemies themselves. This constant change of enemies is in line with the nature of moral panics, which tend to appear unexpectedly, proliferate, and disappear, giving way to other moral panics after a while (Gerő & Sik, 2020). Therefore, researchers have found other enemy images worthy of focusing on throughout the years. As far as we are aware, no one has identified a single enemy whose presence has dominated over the years. Certainly, mention of Brussels and George Soros have been relatively common over a longer period, and they appear in most related studies as cornerstone enemies, while a few others seem to be present only for a limited time. Kenes (2020) notes that Orbán operates with changing enemies; most of these, however, have in common that they involve the 'Western-liberal world'. This enemy-shifting might be a key element in defining an illiberal system, as according to Sata and Karolewski (2022), Orbán's government strongly emphasizes 'friends versus enemies' narratives. Thus, enemy images might be a defining means of supporting such systems. This can be connected to Schmitt's theory concerning the importance of friends and enemies in politics. Pointing out clear enemies can sustain group cohesion and increase support for the stronger group – that is, those considered friends (Schmitt, 1932/2008).



## 7 Conclusion

Our research demonstrates the combined usage of automated text analysis and critical discourse analysis. By utilizing these methods, we were able to qualitatively identify radicalization in the style of the prime-ministerial State of the Nation address speeches and support our observations with keyword frequency-based statistical analysis.

Through the examination of the National Consultations in the period between 2015 and 2022, we were able to identify five main enemy images: Brussels, Gyurcsány, immigrants, Soros, and the opposition. Their presence in the State of the Nation addresses and other keynote speeches of Prime Minister Viktor Orbán was examined. These enemies were present in both types of speeches; however, in accordance with our hypothesis, they exhibited different patterns in terms of frequency. The mentioning of enemy images was less prevalent in the State of the Nation addresses compared to the other keynote speeches, presumably because of the limited, government-close audience. The dominant periods of different enemy depictions, in line with our expectations, occurred in different years. However, their peaks in the State of the Nation addresses did not conform to our hypotheses in several cases, perhaps because these speeches are held in February and dominant enemy images in a given year may only arrive in these speeches with delay.

Attuning the population to fear has long been a political tool (Szabó, 2007). However, Orbán's government has created a system, through the efficient use of national consultations and marketing tools, that puts the overarching emphasis on fearmongering (Sata & Karolewski, 2022). This study has shown that the elements of fear inducement are widely present in the Prime Minister's speeches, which may be due to a strategy of redirecting the focus from certain issues and governing the public discourse in a direction that aligns with Orbán's aims. It is clear that the Prime Minister's enemy images are constantly changing; they come out of the blue, peak, decline, and rise again; some live for very long periods, while others only pop up. In a society that is overly preoccupied with safety (Furedi, 2006), the use of these enemy images and fearmongering tactics can result in substantial political gain. However, regarding the divisive nature of such strategies (Gerő & Sik, 2020), the long-term results can be devastating for society.

## Funding

Zsófia Rakovics's work has contributed to the research that is supported by the ÚNKP-22-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.

The research has contributed to 'The layers of political public sphere in Hungary (2001–2020)' research of ELTE Research Center for Computational Social Science. The referred research project has been supported by the National Research, Development, and Innovation Office (NKFIH), registered with the identifier NKFIH K-134428.

## References

- Altheide, D. L. (2013). Media Logic, Social Control, and Fear: Media Logic, Social Control, and Fear. *Communication Theory*, 23(3), 223–238. <https://doi.org/10.1111/comt.12017>
- Bajomi-Lázár, P. (2013). The Party Colonisation of the Media: The Case of Hungary. *East European Politics and Societies: And Cultures*, 27(1), 69–89. <https://doi.org/10.1177/0888325412465085>

- Boda, Z., & Rakovics, Z. (2022). Orbán Viktor 2010 és 2020 közötti beszédeinek elemzése [Analysis of Viktor Orbán's speeches between 2010 and 2020. Investigating the issue of migration]. *Szociológiai Szemle*, 32(4), 46–69. <https://doi.org/10.51624/SzocSzemle.2022.4.3>
- Cohen, S. (1972). *Folk Devils and Moral Panics: The Creation of Mods and Rockers*. Oxford: Blackwell.
- Cooper, L. (2023). Autocratic Nationalism in Hungary: Viktor Orbán as a Hegemonic Actor. In Kolozova, K. & Milanese, N. (Eds.), *"Illiberal Democracies" in Europe: An Authoritarian Response to the Crisis of Liberalism* (pp. 17–27). The George Washington University.
- Csehi, R., & Zgut, E. (2020). "We won't let Brussels dictate us": Eurosceptic populism in Hungary and Poland. *European Politics and Society*, 22(1), 53–68. <https://doi.org/10.1080/23745118.2020.1717064>
- Demeter, M. (2018). Propaganda against the West in the Heart of Europe. A masked official state campaign in Hungary. *Central European Journal of Communication*, 11(2), 177–197. [https://doi.org/10.19195/1899-5101.11.2\(21\).5](https://doi.org/10.19195/1899-5101.11.2(21).5)
- Fairclough, N. (1995). *Critical Discourse Analysis*. Longman.
- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. Routledge.
- Furedi, F. (2006). *Culture of fear revisited*. A&C Black.
- Gáspár, K. (2021). Demokratikus innovációs eszközök-e a társadalmi konzultációk? [Are popular consultations tools of democratic innovation?]. *Bibó Jogi és Politikatudományi Szemle*, 2021(3), 157–168.
- Gerő, M., Plucienniczak, P. P., Kluknavska, A., Navrátil, J., & Kanellopoulos, K. (2017). Understanding Enemy Images in Central and Eastern European Politics: Towards an Interdisciplinary Approach. *Intersections. East European Journal of Society and Politics*, 3(3), 14–40. <https://doi.org/10.17356/ieejsp.v3i3.365>
- Gerő, M., & Sik, E. (2020). The Moral Panic Button. In E. M. Goździak, I. Main, & B. Suter, *Europe and the Refugee Response* (1st ed., pp. 39–58). Routledge. <https://doi.org/10.4324/9780429279317-4>
- Gerő, M., & Szabó, A. (2017). A társadalom politikai integrációja. A politikai értékcsoportok [The political integration of society. Political value groups]. In Kovách, I. (Ed.), *Társadalmi integráció: az egyenlőtlenségek, az együttműködés, az újraelosztás és a hatalom szerkezete a magyar társadalomban* (pp. 117–154). Belvedere Meridionale.
- Keller, R. (2012). *Doing Discourse Research : An Introduction for Social Scientists*. Sage.
- Kende, A., & Krekó, P. (2020). Xenophobia, prejudice, and right-wing populism in East-Central Europe. *Current Opinion in Behavioral Sciences*, 34, 29–33. <https://doi.org/10.1016/j.cobeha.2019.11.011>
- Kenes, B. (2020). Viktor Orbán: Past to Present. *European Center for Populism Studies (ECPs)*.
- Kopper, Á., Susánszky, P., Tóth, G., & Gerő, M. (2017). Creating Suspicion and Vigilance. *Intersections. East European Journal of Society and Politics*, 3(3), 108–125. <https://doi.org/10.17356/ieejsp.v3i3.366>
- Kopper, A., Szalai, A., & Góra, M. (2023). Populist Foreign Policy in Central and Eastern Europe: Poland, Hungary and the Shock of the Ukraine Crisis. In *Populist Foreign Policy* (pp. 89–116). Palgrave Macmillan, Cham.
- Krekó, P., & Enyedi, Z. (2018). Explaining Eastern Europe: Orbán's Laboratory of Illiberalism. *Journal of Democracy*, 29(3), 39–51. <https://doi.org/10.1353/jod.2018.0043>

- Molnár I. (2020). Válságértelmezés és ellenségképzés a politikai vezetői narratívában. [Crisis interpretation and enemy creation in the narrative of political leadership]. *Politikatudomány Online*, (7), 1–29. <https://www.uni-corvinus.hu/downloads/2svm.13x6szd/politika-tudomany-online-2020-1.8d1.pdf>
- Pócza, K., & Oross, D. (2022). From Deliberation to Pure Mobilisation? The Case of National Consultations in Hungary. *Politics in Central Europe*, 18(1), 79–109. <https://doi.org/10.2478/pce-2022-0004>
- Reményi, P., Glied, V., & Pap, N. (2023). Good and Bad migrants in Hungary. The populist story and the reality in Hungarian migration policy. *Social Policy Issues*, 59(4), 323–344. <https://doi.org/10.31971/pps/162003>
- Sata, R., & Karolewski, I. P. (2022). Caesarean politics in Hungary and Poland. In Agarín, T. (Ed.) *The (Not So) Surprising Longevity of Identity Politics* (pp. 60–79). Routledge.
- Scheppele, K. L. (2022). How Viktor Orbán wins. *Journal of Democracy*, 33(3), 45–61. <https://doi.org/10.1353/jod.2022.0039>
- Schmitt, C. (2008). *The concept of the political: Expanded edition* (G. Schwab, Trans.). University of Chicago Press. (Original work published 1932)
- Sik D. (2019). A negatív integrációtól a félelem szociológiájáig [From negative integration to the sociology of fear]. *Szociológiai szemle*, 29(2), 4–24. <https://doi.org/10.51624/SzocSzemle.2019.2.1>
- Sik, D. (2020). Towards a social theory of fear: A phenomenology of negative integration. *European Journal of Social Theory*, 23(4), 512–531. <https://doi.org/10.1177/1368431019850074>
- Sik, E. (2016). Egy hungarikum: A morális pánikgomb [A ‘Hungaricum’: The moral panic button]. *Mozgó világ*, 42(10), 67–80.
- Spasojević, D. (2019). Transforming Populism – From Protest Vote to Ruling Ideology: The Case of Serbia. In Beširević, V. (Ed.) *New Politics of Decisionism* (pp. 125–140). The Hague: Eleven International.
- Syla, H. (2023). “Keeping Hungary Hungarian”. The Rhetoric of PM Viktor Orbán in the Context of Migration Crises. *ELTE POL-IR Student Paper Series*. 2023/1. 1–33. [https://polir.elte.hu/dstore/document/9043/Hana\\_Syla\\_FINAL\\_to%20be%20uploaded.pdf](https://polir.elte.hu/dstore/document/9043/Hana_Syla_FINAL_to%20be%20uploaded.pdf)
- Szabó, M. (2007). Ellenfél és ellenség a politikában [Opposition and enemy in politics]. *Politikatudományi Szemle*, 16(1), 9–20. [https://real.mtak.hu/112375/1/2007\\_1\\_szabom.pdf](https://real.mtak.hu/112375/1/2007_1_szabom.pdf)
- Szelényi, I., & Csillag, T. (2015). Drifting from liberal democracy. Neo-conservative ideology of managed illiberal democratic capitalism in post-communist Europe. *Intersections. East European Journal of Society and Politics*, 1(1), 18–48. <https://doi.org/10.17356/ieejsp.v1i1.28>
- Tóth, M. Zs. (2015). *Közéleti retorika: Metaforák és mítoszok a magyarországi politikai beszédekben* [Public rhetoric: Metaphors and myths in the Hungarian political speeches]. Eötvös Lóránd University. PhD Thesis.
- Orbán, V. (2010; 2012; 2015; 2016; 2019; 2020; 2022). Annual State of the Nation Addresses. Archive of the speeches of the prime minister of Hungary.
- Walter, A. (2019). Islamophobia in Eastern Europe: Slovakia and the Politics of Hate. *Connections. A Journal for Historians and Area Specialists*. <http://www.connections.clio-online.net/article/id/fda-133254>

Towards a trustful Information Society.  
Comparative analysis of the Czech Republic,  
Hungary, Poland and Romania

---

\* [\[rab.arpad@uni-nke.hu\]](mailto:rab.arpad@uni-nke.hu) (Ludovika University of Public Service, Budapest)

\*\* [\[szikora.tamas@uni-nke.hu\]](mailto:szikora.tamas@uni-nke.hu) (Ludovika University of Public Service, Budapest)

\*\*\* [\[torok.bernat@uni-nke.hu\]](mailto:torok.bernat@uni-nke.hu) (Ludovika University of Public Service, Budapest)

### Abstract

In this article, we present data from a survey on the characteristics of the information society in four Central and Eastern European countries. The 25-minute representative telephone survey was conducted in Hungary, Romania, Poland and the Czech Republic. The research covers internet usage habits, actors of online trust, fear of manipulation and news consumption habits, and last but not least, general attitudes towards the internet. The four databases are available in English and the four questionnaires are available in the language of the country in which they are annexed to the article. The main variables of the database are presented in the article.

**Keywords:** online trust, awareness, internet usage, Czech Republic, Hungary, Poland, Romania

## 1 Introduction

The spread of the information society and digital culture has changed our world. New forms of communication and cooperation have emerged, and our social subsystems are fundamentally changing. The rapid development of the information society in each of the countries studied is undeniable, but of course important differences can also be observed. In order to investigate the perceptions of a society in the online space, we needed to have something to compare them with. To this end, we compared four European information societies with many similarities but also several differences. Comparable research into the digital culture of the four countries on the same platform does not exist in this depth. The four countries were selected by comparing Eurostat data and reviewing the most important academic literature on them. (European Bank, 2020; Mirke, Kasparova, & Cakula, 2019, European Commission, 2019; Statistical Office of Poland, 2020; Szarek-Iwaniuk & Senetra, 2020).

Online platforms and communication services are now at the centre of public discourse, and for good reasons. First, on a rather descriptive level, these services have become not just widespread, but an indispensable part of our personal and social life. People

---

spend more and more time in the online world, because they increasingly manage their lives through digital services. Second, on a more analytical level, the new infocommunication technologies have substantially changed some of the central elements of our everyday lives: how we contact each other, how we communicate in private and in public, how we work, how we travel and how we get information. While it is important to track the trends of internet use in different generations and in society as a whole, it is also becoming crucial to look beyond the patterns of use, and to attempt to understand people's attitude towards these services. Our research focuses on the latter and identifies key elements of this attitude: the extent to which people trust internet services, how consciously and prudently they navigate online, and how aware they are of possible risks. Trust, awareness, and alarm: these features are key not only to gain a deeper understanding of societies running online, but also to identify the reasons why regulators bear responsibility in the formation of the digital world.

Hungarians are the most afraid, among the populations surveyed, of online communication having a negative effect on their personal relationships. According to 60.5% of the total Hungarian population, online communication degrades the quality of personal relationships. This view was held by 50.7% in the Czech Republic, 38.9% in Poland and only 29% in Romania. At the same time, the region is uniform in its judgement on the benefits of the Internet. The evaluation of new communication opportunities and the positive impact of digital culture on society did not differ significantly from country to country.

Facebook has a remarkably strong market position among social media platforms, and its position is the most dominant in Hungary. It is typical for the whole region that, compared to Facebook, the other social media interfaces hardly get a kick at the ball, and of the four countries, this is the most striking in Hungary. While in the Czech Republic and Poland an age gap can be identified in the use of Facebook, in Hungary and Romania the older age groups are also definitely present on the platform.

A high degree of awareness about internet use is not typical in any of the countries examined. Although people are aware of the dangers to be found online, they do not defend themselves or their families from them through training or filtering software. The same is true regarding data protection: the majority of the respondents were aware of the importance of data protection and is distrustful of online data controllers, yet they do not do anything about the security of their data.

Significant differences can be observed between the countries of the region in terms of trust in information sources and fear of manipulation. A comparative analysis of the four countries revealed significant differences with regard to fear of manipulation and trust in the information received. The vast majority of the Hungarian population is not afraid of being manipulated by specific actors in the online world. Hungarian behaviour in this regard is most similar to that of Romanian users. In Hungary, people are most distrustful of advertisers (23%) and in Romania of politicians (22.5%). The general fear of manipulation is higher in the Czech Republic and Poland, and the highest levels of mistrust (measured for advertisers in both countries) were also much higher (52% and 40%, respectively).

There is significant potential for development in the field of e-business and e-government among the older residents of the region. Information literacy is already satisfactory,

but there is still a long way to go in terms of levels of usage, which represents great potential for the public and business sectors in all four countries. There are also great opportunities in the field of quality-of-life improvement and preventive online services.

## 2 Methods

This study was conducted by the Institute of the Information Society of the Ludovika University of Public Service and covered four Central European countries: the Czech Republic, Hungary, Poland and Romania, with the aim of examining the characteristics of the use of information technology by the adult population in the region, mainly for communication purposes.

A telephone survey was conducted in October and November 2019, and the results are representative of the population over 18 years of age in the four countries, categorized by age, gender, education, type of settlement and region.

The questionnaire consists of 50 questions and was conducted by telephone interviews lasting twenty-five minutes. The four databases are fully comparable with each other and have been weighted for this purpose, the weight variable being found in the databases.

## 3 Description of data and variables

There are eight files attached: four data files in SPSS format (Hungary, Romania, Czech Republic, Poland) and four questionnaires in the original languages (DOI:10.5281/zenodo.8267401)

The most important variables in the data set are:

### Dimensions

TELTIP – Type of settlement

D2\_1 – Age in years

D2 – Age categories

D3 – Gender of respondent

D4 – Education – 6 categories

D5\_1 – Persons in households

D6\_1 – Children under 14 years

E1 – Do you have a computer at home?

E2 – Do you have a smartphone?

E3 – Do you use the internet wherever you have access to it (at home, at an internet café, etc.)?

E4 – What tools do you use to access the internet? (PC, laptop, tablet, smartphone)

E5 – Do you have a home internet subscription?

E6 – Do you have a mobile internet subscription?

E7 – Do you use any filtering software to protect your child/children from harmful content?

I1 – How often do you use the internet?



I2 – How often do you use the internet for the following purposes? (E-mail, chat, phone, Facebook, partner search, news, getting information for work, getting information for learning, online banking, office administration, streaming music, streaming films, online gaming, online shopping).

### **Manipulating**

I3 – how afraid are you of being manipulated by them? (Social platforms, search engines, advertisers, journalists, politicians, government, civil organizations, experts, foreign powers, influencers, friends, large companies, content produced by machines)

M3 – How much confidence do you have in the credibility of the information you receive? (Friends, TV, radio, print, social media, online news)

### **Internet usage**

I4 – Have you ever refrained from using the internet because you were worried that everything in the digital world would be left behind?

I5 – Have you ever paid for any extra facilities or premium services linked to a basically free online service? For example, Spotify or YouTube Music ad-free, Tinder extra features, Twitch subscription, premium article reading, etc.

I7 – Have you ever been harassed on the internet?

I8 – How many friends or partners do you chat with on a more or less regular basis?

K1 – How would you describe your own communication activities on social media portals?

K2 – Have you ever downloaded or shared an article or video that you knew was posted on the internet without permission, in an illegal way? What was your reason for doing so?

K4 – Have you produced or shared content online that you later regretted?

K6 – Do you share memes?

U1 – Are you using a government client gateway?

H3 – How do you watch films online?

H4 – How do you listen to music on the internet?

H5 – Do you know someone you met online and have kept in touch with in person?

### **Facebook**

F1 – How often do you post on Facebook?

F2 – How often do you read news on Facebook?

F3 – Have you ever had a post deleted by Facebook or been temporarily banned from the service?

F4 – Were you informed of the reason for the deletion and/or blocking?

F5 – Did you request a change in the decision to delete and/or block?

F6 – What was the outcome of the application?

F7 – Have you ever adjusted your privacy settings on Facebook, eg. who can see your posts?

### **News reading**

H1 – How would you describe your news consumption in terms of the following content? (Public news, boulevard news, sport news, scientific news, general news)

H2 – Are you interested in public positions that differ from your convictions? Which of the following answers best suits you?

M2 – How often do you use the following information sources to find out about world news? (Friends, television, radio, print, social media, online news site)

### **Attitudes**

M4 – In your opinion, do you think that online/digital communication opportunities and social portals enhance or diminish the quality of personal relationships?

M5 – Get a lot of information about your contacts easily and at once, easy to contact people you know and get feedback quickly, There is a fear that direct/face-to-face contact will be replaced by online contact, Too much time taken up by online communication, using social networking sites

M6 – the internet is advantageous because I can get important information quickly and directly b) the internet is risky because I can get uncontrolled information, the internet is beneficial because ordinary people can share their views with the social public b) the internet is risky because views that are dangerous to social coexistence can be freely disseminated, I like public news and events to reach me directly b) I like public events to reach me with analysis and evaluation, the activity of newspapers and journalists is a very important part of the social dialogue b) in the age of the internet, the need for newspapers and journalists is decreasing, anonymous expression of opinion/commenting on the internet is a good thing, because it allows many people to express their opinions freely b) anonymous expression of opinion/commenting on the internet is harmful, because it encourages irresponsible statements

M7 – Overall, do you see the rapid development of digital services and online activities as an opportunity or a risk for individuals and society?

### **Data awareness**

M14 – If you were offered a very attractive free service, which of the following details would you provide in order to qualify? (Address, tax number, age, plate number, bank account number, bank card number, information about traveling, information about health)

M15 – How secure do you feel this personal data is in the following areas? (Public sector, Facebook, Google, Your own smartphone operating system, On your own computer, With your account-holding bank, Internet advertising portal, Your telecoms provider, A company/portal offering online services (e.g. food delivery, hosting, travel arrangements, etc.), medical office, car service, local government).

### **Competing interests**

The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

### **Ethics Statement**

The source questionnaire and data collection projects are ethically approved by the ESS ERIC Research Ethics Committee. Informed consent from respondents was obtained (European Social Survey, n.d.).

## References

- Andersen, K. & Strömbäck, J. (2021). Media Platforms and Political Learning: The Democratic Challenge of News Consumption on Computers and Mobile Devices. *International Journal of Communication*, 15, 300–319. <https://ijoc.org/index.php/ijoc/article/download/15511/3320>
- Czech Statistical Office (2020) *Information society in figures 2020 – Czech Republic and EU*. Publication code: 061005-20. Czech Statistical Office 21 Apr 2020. <https://csu.gov.cz/produkty/information-society-in-figures-2020> Accessed: 27-11-2024.
- Dubois, E. & Blank, G. (2018). The echo chamber is overstated: the moderating effect of political interest and diverse media. *Information, Communication & Society*, 21(5), 729–745. <https://doi.org/10.1080/1369118X.2018.1428656>
- European Bank for Reconstruction and Development (2020) *Romania Country Strategy 2020–2025*. <https://www.ebrd.com/strategy-and-policy-coordination/strategy-for-romania.pdf> Accessed: 27-11-2024.
- European Commission (2019) *Digital Government Factsheet 2019 Romania*. [https://interoperable-europe.ec.europa.eu/sites/default/files/inline-files/Digital\\_Government\\_Factsheets\\_Romania\\_2019.pdf](https://interoperable-europe.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_Romania_2019.pdf) Accessed: 27-11-2024.
- Graf, J., Erba, J. & Harn, R.-W. (2017). The role of civility and anonymity on perceptions of online comments. *Mass Communication and Society*, 20(4), 526–549. <https://doi.org/10.1080/15205436.2016.1274763>
- GUS (2020) *Information society in Poland in 2020*. GUS Statistical Office in Szczecin. <https://stat.gov.pl/en/topics/science-and-technology/information-society/information-society-in-poland-in-2020,1,7.html> Accessed: 27-11-2024.
- Haider, J. & Sundin, O. (2019). Information literacy challenges in digital culture: conflicting engagements of trust and doubt. *Information, Communication & Society*, 25(8), 1176–1191. <https://doi.org/10.1080/1369118X.2020.1851389>
- Helberger, N. (2019). On the Democratic Role of News Recommenders. *Digital Journalism*, 7(8), 993–1012. <https://doi.org/10.1080/21670811.2019.1623700>
- Leaning, M. (2019). An Approach to Digital Literacy through the Integration of Media and Information Literacy. *Media and Communication*, 7(2), 4–13. <https://doi.org/10.17645/mac.v7i2.1931>
- Liu, X. & Lu, J. (2020). Does the Internet Erode Trust in Media? A Comparative Study of 46 Countries. *International Journal of Communication*, 14, 5822–5837. <https://ijoc.org/index.php/ijoc/article/viewFile/14552/3277>
- Mirke, E., Kašparová, E. & Cakula, S. (2019) Adults' readiness for online learning in the Czech Republic and Latvia (digital competence as a result of ICT education policy and information society development strategy). *Periodicals of Engineering and Natural Sciences*, 7(1), 205–215.
- Osborne, C. L. (2018). Programming to Promote Information Literacy in the Era of Fake News. *International Journal of Legal Information*, 46(2), 101–109. <https://doi.org/10.1017/jli.2018.21>

- Padyab, A., Päivärinta, T., Ståhlbröst, A. & Bergvall-Kåreborn, B. (2019). Awareness of Indirect Information Disclosure on Social Network Sites. *Social Media & Society*, 5(2), 1–14. <https://doi.org/10.1177/2056305118824199>
- Riedl, M. J., Whipple, K. N. & Wallace, R. (2021). Antecedents of support for social media content moderation and platform regulation: the role of presumed effects on self and others. *Information, Communication & Society*, 25(11), 1632–1649. <https://doi.org/10.1080/1369118X.2021.1874040>
- Russmann, U. & Hess, A. (2020). News Consumption and Trust in Online and Social Media: An In-depth Qualitative Study of Young Adults in Austria. *International Journal of Communication*, 14, 3184–3201. <https://ijoc.org/index.php/ijoc/article/download/13774/3115>
- Stevic, A., Schmuck, D., Karsay, K. & Matthes, J. (2021). Are Smartphones Enhancing or Displacing Face-to-Face Communication With Close Ties? A Panel Study Among Adults. *International Journal of Communication*, 15, 792–813. <https://ijoc.org/index.php/ijoc/article/download/14796/3351>
- Thorson, K., Cotter, K., Medeiros, M. & Pak, C. (2019). Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society*, 24(2), 183–200. <https://doi.org/10.1080/1369118X.2019.1642934>
- Vraga, E. K. & Tully, M. (2021). News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication & Society*, 24(2), 150–166. <https://doi.org/10.1080/1369118X.2019.1637445>

---

## BOOK REVIEW

Singh, J. (2023). *Natural Language Processing in the Real World: Text Processing, Analytics, and Classification*. CRC Press, Taylor & Francis Group

---

<https://doi.org/10.17356/ieejsp.v1i1.1421>

*Natural Language Processing (NLP) in the Real World* presents a structured guide to the fundamentals of NLP, combining theoretical concepts with practical applications. Covering essential techniques like text preprocessing, sentiment analysis, text classification, and named entity recognition (NER), the author emphasizes hands-on Python implementations, making it accessible to beginners and professionals alike. The book's strength lies in its focus on real-world problems, such as social media analysis and customer feedback. NLP has emerged as an indispensable tool for analysing large amounts of text data, especially with the exponential growth of unstructured information across industries like e-commerce, social media, healthcare, and finance. The book comprehensively explores NLP techniques with a focus on real-world applications. It aims to bridge the gap between theoretical understanding and practical implementation of NLP tasks (Simske & Vans, 2021). This review critically assesses the book's chapter-wise content, evaluating the strengths and weaknesses of its methodology, as well as its contributions to social science research, data science and academia. The analysis also highlights the book's limitations, especially in addressing the complexities of modern NLP research.

Natural Language Processing (NLP) is a branch of artificial intelligence that focuses on the interaction between computers and human languages by transferring it to the algorithm (Dyshel & Lane, 2023). It involves text processing, sentiment analysis, machine translation, speech recognition, and text classification. By using techniques like tokenization, stemming, and machine learning, NLP enables machines to understand, interpret, and generate human language. Modern applications include chatbots, virtual assistants, language models (e.g. GPT), and more. NLP is critical in various domains, including social media analysis, healthcare, and customer service automation. Through different sections and chapters, the author explains important NLP concepts, ideas, data curation, data processing, implementation of NLP application, and implementation of NLP application in the real world.

The book deals with questions such as how can we preprocess and prepare text data for machine learning models, what are the most effective methods for performing sentiment analysis on textual data, which classification algorithms can be used to categorize

---

text, how can Named Entity Recognition (NER) be applied to identify specific entities within text (Fortino, 2021), and finally, what are the practical challenges in applying NLP to real-world scenarios like social media analysis or customer feedback? These questions guided the author to articulate the book systematically.

*Natural Language Processing in the Real World* explains how to clean, tokenize, and normalize text using methods like stemming, lemmatization, and stop-word removal, making it ready for machine learning. It demonstrates how machine learning algorithms, such as Naive Bayes or Support Vector Machines (SVM), classify text sentiment, social media and product reviews (Atkinson-Abutridy, 2022). It explains how to identify and classify entities like the names of people, organizations, or places using NER algorithms, employing libraries like spaCy. Most importantly, the book highlights challenges in applying NLP to real-world contexts, such as noisy text data, social media, and business-related use cases, and provides step-by-step Python code to overcome these issues.

The book has six sections covering twelve chapters to explain NLP as its central idea. The introductory discussion in Section I (Chapter 1) introduces fundamental NLP concepts in an accessible manner. While it covers standard material like tokenization, part-of-speech tagging, stemming, and lemmatization, it offers little to experts in the field (Iezzi et al., 2020). It could benefit from broader coverage of recent innovations like transformer-based tokenization or advanced embeddings.

Section II (Chapter 2) discusses real-world data challenges, especially in environments where pre-labelled datasets are scarce. It allows us to learn how to identify, access, and extract data from diverse sources. The author's experience with industry datasets adds depth here, but it could benefit from more exploration of cutting-edge scraping techniques and API integration methods. Greater attention to data ethics, especially in large-scale extraction, could enhance the strength of this section.

Section III (Chapter 3) gives substantial coverage of key preprocessing steps, including dealing with unstructured text and feature extraction. The author provides practical Python code examples, which enhance the book's usability. The chapter thoroughly covers text cleaning, tokenization, and vectorization, offering clear Python implementations, however, it could offer more advanced techniques like sentence embeddings or multilingual text preprocessing. This section permits learners to capture master techniques for cleaning and transforming raw text data into structured formats for analysis (Iezzi et al., 2020), and to understand feature extraction and vectorization methods.

Chapter 4 of the section is one of the book's highlights. The author seamlessly transitions into various NLP models, from traditional ones like TF-IDF (Term Frequency-Inverse Document Frequency) to more advanced machine learning algorithms (Dyshel & Lane, 2023). The balance between theoretical explanation and application is well-maintained, though a more detailed discussion on transformer-based models would have added value, given their dominance in the field (Dyshel & Lane, 2023). The model diversity here is commendable, ranging from traditional machine learning to newer deep learning models, ensuring a broad overview. More focus on state-of-the-art models such as transformers could be beneficial, as these are now industry-standard in NLP tasks. Nevertheless, the omission of detailed hyperparameter tuning or model optimization strategies leaves gaps for more advanced readers.



Section IV (Chapters 5 & 6) focuses on real-world use cases of NLP, emphasizing industry applications. The author provides rich examples across different domains, making the content particularly valuable for professionals in e-commerce, finance, and customer service. However, the sections are somewhat disjointed, lacking a unifying narrative between the case studies. In Chapter 5, the real-world applications in various industries are well-covered, giving readers clear insights into NLP's practical impacts in sectors like finance and healthcare. It indicates that better integrating practical examples with theoretical insights could strengthen the learning experience. Chapter 6 highlights emerging areas of NLP, such as legal text mining and healthcare applications. Though the range of applications is limited, the Chapter could explore how emerging NLP applications evolve in response to large language models.

Section V (Chapters 7 & 8) presents information extraction techniques, covering both rule-based and machine-learning approaches. However, the book could benefit from deeper coverage of neural network-based extraction models, such as BERT and GPT-based frameworks, which are now becoming industry standards. In Chapter 7, the author offers solid coverage of entity extraction and summarization techniques, providing practical examples. This Chapter is an excellent source to gain knowledge of information extraction techniques and text summarization (Ignatow & Mihalcea, 2018). Also, it explores models for entity recognition and transformation. However, while practical, it lacks depth in explaining sophisticated extraction models, particularly those utilizing pre-trained models. Advanced techniques such as abstractive summarization models like PEGASUS are under-explored.

In Chapter 8, the author explains various text categorization algorithms, from Naïve Bayes to Support Vector Machines and the use of affinity models. The inclusion of performance metrics and error analysis for each model is a commendable touch. However, the absence of more recent advances like zero-shot classification is noticeable, which is meticulously relevant in today's fast-evolving NLP landscape (Wang et al., 2023; Yu et al., 2023). However, excellent explanations of classification techniques, from logistic regression to support vector machines, with relevant code examples are promising. However, the challenge lies in balancing foundational theory with cutting-edge developments to meet a broad audience's needs. Nonetheless, the section is equipped enough to develop skills in text categorization techniques among learners who can learn to use models like Naïve Bayes and Support Vector Machines for classification tasks.

Section VI comprises four chapters (9-12) and delves into advanced NLP applications, which marks an essential shift from foundational concepts to real-world solutions. Each chapter focuses on practical implementation, moving beyond theoretical discussions and offering insights into how NLP can be applied to develop functional, scalable systems in real-world environments. Chapter 9 focuses on one of the most visible and widespread applications of NLP – chatbots. The author explains the basic principles of chatbot development, including rule-based approaches and machine learning-driven solutions. The chapter provides step-by-step guidance, with Python code and tools such as NLTK and spaCy, for building a basic chatbot. The real-world focus on chatbots ties in well with the book's overarching theme of practical NLP. The explanation of both rule-based and machine learning methods allows readers to explore multiple development paths, based on their

needs or technical expertise. However, the chapter's limitation lies in its exclusion of more advanced conversational AI models, such as transformers (e.g. GPT-3), which dominate modern chatbot development. The absence of neural-based approaches could leave readers wanting more depth in this rapidly evolving area.

Chapter 10 provides an in-depth look at how NLP can be used to analyse customer feedback from platforms like e-commerce sites and social media. It focuses on extracting actionable business insights through sentiment analysis, topic modelling, and aspect-based sentiment analysis, helping businesses improve products and enhance user experiences (Sailunaz et al., 2018; Seyeditabari et al., 2018). It includes practical Python examples, making the concepts accessible to business analysts, data scientists, and product managers. The Chapter primarily relies on traditional machine learning models, such as Naïve Bayes and SVM, for sentiment classification and feature extraction, offering a comprehensive guide to these techniques. While this makes it suitable for those looking to apply NLP to real-world problems quickly, it lacks coverage of advanced deep learning methods, such as transformers and BERT, which are now widely used in NLP tasks. That limits the chapter's relevance for readers seeking more cutting-edge techniques. The author also addresses challenges in handling noisy, unstructured text by explaining effective preprocessing methods like tokenization and stop-word removal (Atkinson-Abutridy, 2022). The chapter provides valuable strategies for identifying sentiment polarity and extracting specific product features mentioned in reviews. However, it offers only a superficial treatment of more complex issues, such as sarcasm detection and opinion spam, which are critical in customer review analysis but not fully explored.

Chapter 11 focuses on integrating NLP into recommendation systems, a vital tool in e-commerce and media platforms. Singh demonstrates how sentiment analysis and topic modelling can predict user preferences by analysing text-based interactions, product descriptions, and reviews. The chapter's strength lies in its practical approach, offering step-by-step guides with Python libraries like NLTK and spaCy, making it accessible for professionals. However, it has limitations, particularly its lack of discussion on advanced deep learning techniques such as neural networks and transformers, which limits its relevance for cutting-edge research. It also misses a critical discussion on ethical issues like bias and over-personalization in recommendation systems. Despite these gaps, the author addresses challenges like noisy text data and the user cold start problem, providing practical solutions.

The last chapter broadly examines how NLP can enhance business intelligence (BI). It shows how textual data from various sources – such as emails, customer reviews, or news articles – can be processed and analysed to deliver insights that support business decision-making. This chapter highlights one of the most crucial applications of NLP – turning unstructured text into actionable data for businesses. The author effectively showcases how BI tools can integrate NLP and provides relevant case studies to show its real-world impact. The coverage of this chapter is somewhat generic, focusing more on traditional NLP techniques. It overlooks a discussion on how emerging trends like deep learning, transformers, and knowledge graphs can enhance business intelligence. This omission reduces the relevance of the chapter for businesses seeking more cutting-edge solutions.

Section VI of the book highlights practical NLP applications like chatbots and recommendation systems, providing accessible explanations and Python code for implementation. However, it mainly focuses on traditional methods and neglects advanced models like transformers, limiting its demand for researchers. Though there is a lack of discussion on ethical implications, the section offers useful insights for practitioners. Its limitations may hinder its relevance for advanced research.

The methodology of the book is practical, emphasizing traditional machine learning techniques like Naive Bayes and SVM, with essential NLP preprocessing steps such as tokenization and stemming. The accessible, Python-driven approach is ideal for real-world applications like sentiment analysis. While the methodology is more suited for practical implementation and introductory learning, the book provides a solid foundation for applying NLP techniques in real-world settings, contributing to academic learning and practical applications across fields. However, it lacks depth in advanced methods like deep learning, limiting its use for cutting-edge research. Additionally, the book does not adequately address ethical concerns in NLP applications, such as bias or more complex issues like sarcasm detection and opinion spam.

The volume is particularly recommended for professionals in fast-paced industries as it offers practical, code-based solutions for immediate implementation. Its interdisciplinary approach appeals to a wide audience, including data scientists and social researchers, with accessible explanations and coding examples suited for beginners. Covering sentiment analysis, information extraction, and text categorization, it provides actionable skills applicable to business, technology, and social science research, especially for analysing large-scale textual data such as social media posts or surveys (Grimmer et al., 2022). It addresses common challenges in text processing, such as noisy data, context understanding, bias in models, and language diversity. The author offers practical solutions and applies NLP in real-world contexts such as sentiment analysis and customer reviews. Although the book briefly touches on bias mitigation, it primarily focuses on providing scalable, real-time Python-based solutions for large-scale data, making it especially valuable for social researchers tackling complex societal issues using NLP techniques. Though it could explore more advanced techniques like transformer models, its focus on real-world applications ensures it remains a useful resource for those looking to apply NLP in industry or research settings.

MOHAMMAD ASHRAFUL ALAM<sup>1,2,\*</sup>

<sup>1</sup> ELTE Eotvos Lorand University, Doctoral School of Sociology, Faculty of Social Sciences, Budapest-1117, Hungary.

<sup>2</sup> Department of Criminology and Police Science, Mawlana Bhashani Science and Technology University (MBSTU), Tangail-1902, Bangladesh.

\* Correspondence: maalam@student.elte.hu; maalam.cps@gmail.com; Orcid: 0009-0005-7064

## References

- Atkinson-Abutridy, J. (2022). *Text analytics: An introduction to the science and applications of unstructured information analysis*. Chapman and Hall/CRC. <https://doi.org/10.1201/9781003280996>
- Dyshel, M., & Lane, H. (2023). *Natural Language Processing in Action* (2nd, Version 8 ed.). Manning Publications. <https://www.manning.com/books/natural-language-processing-in-action-second-edition>
- Fortino, A. (2021). *Text Analytics for Business Decisions: A Case Study Approach*. Mercury Learning and Information.
- Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton University Press.
- Iezzi, D. F., Mayaffre, D., & Misuraca, M. (Eds.). (2020). *Text Analytics: Advances and Challenges (Studies in Classification, Data Analysis, and Knowledge Organization)* (1st ed.). Springer.
- Ignatow, G., & Mihalcea, R. (2018). *An Introduction to Text Mining: Research Design, Data Collection, and Analysis*. Sage Publications.
- Sailunaz, K., Dhaliwal, M., Rokne, J., & Alhaji, R. (2018). Emotion detection from text and speech: a survey. *Social Network Analysis and Mining*, 8, 28. <https://doi.org/10.1007/s13278-018-0505-2>
- Seyeditabari, A., Tabari, N., & Zadrozny, W. (2018). Emotion detection in text: a review. *ArXiv Preprint ArXiv:1806.00674*. <https://doi.org/10.48550/arXiv.1806.00674>
- Simske, S., & Vans, M. (2021). *Functional Applications of Text Analytics Systems*. River Publishers.
- Wang, Y., Feng, L., Song, X., Xu, D., & Zhai, Y. (2023). Zero-Shot Image Classification Method Based on Attention Mechanism and Semantic Information Fusion. *Sensors*, 23(4), 2311. <https://doi.org/10.3390/s23042311>
- Yu, Y., Zhuang, Y., Zhang, R., Meng, Y., Shen, J., & Zhang, C. (2023). ReGen: Zero-Shot Text Classification via Training Data Generation with Progressive Dense Retrieval. In A. Rogers, J. Boyd-Graber, & N. Ozaki (Eds.) *Findings of the Association for Computational Linguistics: ACL 2023* (pp. 11782–11805). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.findings-acl.748>