Doctoral School of Earth Sciences

Spatiotemporal and multilingual Semantic Machine Learning Analysis of Social Media Data for the recent protests in Europe
– based on Twitter data –

Tamás Kovács

Supervisors:

Dr. Ádám Németh

Prof. Dr. Klára Siposné Kecskeméthy

University of Pécs

Faculty of Sciences

Pécs, 2022

Table of Content

List of figures

List of tables

INTRODUCTION

The continuously increasing digital data produced either by individuals or by computers is changing the views of various sciences, forming our knowledge of humankind and nature. Moreover, since the beginning of the 21st century, the rapid spread and increasing popularity of social media has generated a valuable data source for analyzing or even predicting human behavior based on the regular activities of individuals. The context analysis of social media pitches a series of questions due to its complexity in understanding human behavior. Therefore, one of the central debates concentrates on the basis and explanations of the patterns for social media analysis. When approaching social media for a social movement or protest analysis, the question appears as to what information could help understand these movements.

The examination of information from social networking sites can be instrumental in achieving a thorough interpretation of the human environment and social dynamics. Such research projects usually utilize a so-called passive crowdsourcing approach, where data is generated collectively by users of a social media platform, but without direct contributions to a specific research or crowdsourcing project, as opposed to traditional (or active) crowdsourcing such as OpenStreetMap (Connors et al., 2012, Muller et al., 2015). The emerging field of passive (or opportunistic) crowdsourcing relies on such data, granting empirical investigations that usually build on a semi-automated data collection process. The users either provide data in the form of text (e.g., Twitter) or images (e.g., Instagram), which are often paired with sensor-obtained information (e.g., location) and uploaded online (Peña-López et al., 2014, Huang et al., 2020). Despite the unique usability of this research field, relatively few papers published so far have concentrated on multi-dimensional analysis applications of social media data in the analysis of collective actions, such as protests.

It is no exaggeration to say that everyone these days coordinates and disseminates information through the internationally distributed social network has new chances and opportunities (Burgess et al., 2012). Through web-based social networks that transcend physical boundaries, large numbers of like-minded individuals share and discuss their local or global problems (Khondker, 2011). Social media platforms have become standard communication tools that enable people to connect by sharing opinions and debating social issues (Kim et al., 2013). Furthermore, social media may be an effective device or certain cases, a capable and complex weapon for protesters in the context of social movements. It can be used to broadcast individual ideas that urge others to take collective action, thus advancing protest mobilization. (Choi and Park, 2015).

Protests and social movements appear from time to time in various locations. They depend on many factors, bear various characteristics, and they could reveal an extended strategic complexness and relations with other networks. In historical research, these aspects are predominantly referred to as protest generators, trigger events, or attractors (Cantor, 2021). Examples of such characteristics include events that commonly engage large groups, such as markets. It is enough to consider the French Revolution that ended with the French Consulate's formation in 1799. According to our current knowledge, lack of food and increased prices ignited an outbreak of widespread anger in the villages and towns of Paris (Furet, 1981). As Voltaire wrote a few years before the events, Parisians required nothing besides comic opera but some „white bread," referring to a general problem in French society. Indeed, more than three hundred riots and excursions to pillage grain were documented in over three weeks locally (Kaplan, 1985).

The French Revolution is often considered an origo that transformed the European political culture for over a century. For example, under the influence of reading about the history of the Revolution, Sándor Petőfi, poet and liberal revolutionary, became one of the key figures of the Hungarian Revolution of 1848, fighting for independence from Austria (Hermann & Bona, 1996). These days, our interconnected world can establish and propagate local events globally and connect distant but similar interests in the online sphere. The first Occupy demonstration was established in New York City's Zuccotti Park on 17 September 2011, motivated by the same interest as the French Revolution, namely social and economic inequality. However, within a few weeks, it spread across the online sphere, forming local but interconnecting protests in over 951 cities across 82 countries worldwide (Gleason, 2013).

The use of digital media to organize, communicate and coordinate protest activities between dispersed national and international people, organizations and associations was a typical underlying pattern of these mobilizations. Democracy and mass communication scholars are increasingly emphasizing the role of information technologies in facilitating democratization and fostering transition in young and developing democracies. Moreover, it consolidates and stabilizes engagement and enhances participation in established democracies (Lynch, 2011; Shirky, 2011).

Contemporary scholarship has grown increasingly interested in the role of digital media in protest events; studies have asserted that the internet can assist activists in diversifying their engagement repertoires, expanding their engagement beyond previous spatial and temporal constraints, and organizing and coordinating protest participation more effectively (Bennett, 2003; Van de Donk, Loader, Nixon, & Rucht, 2004, Earl & Kimport, 2011; Castells, 2015). In addition, recent scholarly work from a variety of methodological and disciplinary perspectives (Bennett & Segerberg, 2012; Earl et al., 2013; Theocharis, 2013; Jungherr & Jürgens, 2014) along with popular accounts has emphasized the capacity of social media to assist activists in more effectively managing the complexities of mass protest organization and coordination. For instance, the existing literature discusses the numerous benefits of microblogs for social movements and activists (Shirky, 2009, Earl & Kimport, 2011), while many qualitative empirical studies examined the function of social media between activist societies in the Occupy movements in the United States and the Netherlands (Mercea, 2013, Penney & Dadas, 2014). However, few studies have conducted empirical evaluations of the evolutions in protest dynamics that have occurred as a result of social media use, and comparative evaluations that could consolidate knowledge obtained from individual case studies are still absent (see Bennet & Segerberg, 2013; Earl et al., 2013; González-Bailón et al., 2013).

This work aims to reveal recent scholarship limitations and offer a unique analytical methodology to solve emerging questions. This dissertation predominantly concentrates on the preliminary stage of protest activity, its spatiality, and interconnectivity. It will suggest that spatial distribution patterns based on individuals' social media posts and their reaction time are similar in countries with similar political systems and interests. This idea is partly based on a well-known political sociology theory, Political Opportunity Theory (Meyer & Minkoff, 2004), that relies on the presence or the absence of a specific political opportunity. It argues further that the actions of the activists are dependent on these. The primary benefit of the theory is that it describes why social movements increase their activity at a given time. In addition, It will also be discussed the incorporation of semantic and location information from Twitter data as dynamic features for protest analysis and short-term spatiotemporal protest prediction.

Attribute impacts extracted from massive social media analysis offer excellent potential for protest analysis and prediction. First, however, it is crucial to understand the pros and cons of selecting representative data. This scholarly work aims to discover and assess spatial relationships between protest events and nearby social media activities and to assess the potential influence of social media posts on protest prediction models. Although current protest prediction models show increasing accuracy, little emphasis has been placed on understanding how and why predictive features behave the way they do in space and time.

LITERATURE REVIEW

This chapter reviews the related literature on our subject of study, starting with the theoretical theses, which were adopted by individual case studies. This will help create a solid base for comprehending the topic's actuality and contribution to the research field. Let us start immediately with a limitation. Although there exists pervasive literature on the topic of various types of social movements according to their nature, purpose, claim, or even their persistence, scope, or form of action, the present work will primarily focus on revolutionary movements that are trying to surpass the political, economic, or social limits to implement a new social and political order. As Charles Tilly described it as "a social movement advancing exclusive competing claims to control of the state, or some segment of it" (Tilly, 1995). In Diani's consideration, a social movement is an interaction network between a group of individuals or groups engaged in political or cultural conflict based on shared collective identities (Diani, 1992).

1 The collective action

The essential building blocks of every social movement are a collective action and, within it, the people and their motivation emerging from their feelings. Collective action refers to an action taken jointly by a group of people whose goal is to improve their condition and achieve a common goal (Marshall 1998). By nature, it could manifest itself in various forms, from irregular and cataclysmic forms such as protests, crowds, or pressure groups. Some of these affect physically present people (e.g., protests), while others may affect very far apart people, such as rumors, even though the second may feed the first.

As Tilly framed, five significant components consist of collective action: interest, organization, mobilization, opportunity, and collective action. The interest is the sum of the gains and losses resulting from the interaction of groups. The organization is that aspect of a group's structure that most directly impacts its ability to operate on its interests. Mobilization is how a group gains coordinated control over the resources (e.g., goods, votes, workforce) required for its action. The analysis of mobilization deals with the ways that groups obtain resources and use them for collective action. Finally, the opportunity is the relationship between a group and its world where the collective action results from altering interests, organization, mobilization, and opportunity combinations (Tilly, 1978).

The research field of collective action has undergone significant change since its establishment in the 20th century. Although the classical theory of crowd psychology already regarded emotions as triggers of collective behavior up to the '60s (Le Bon & Merton, 1960, Krosigk & Sighele, 2008), the subsequent decades based on irrational behavior have gradually been replaced by an emphasis on the rationality of collective action (Olson, 1971). Nevertheless, in analyzing social movements, individuals function as social change agents. The focus compass should point to those social behaviors without which social change would not occur and, therefore, necessary for social mobilization. Therefore, social movement theories have attempted to comprehend what gives rise to social movements and why some individuals are more likely than others to participate in social mobilizations (Davies, 1962; Gurr, 2015).

1.1 Theories and changes in the political organization and coordination

Franklin Henry Giddings coined the term „collective behavior," which was subsequently used by Robert Park and Ernest Burgess, Herbert Blumer, Ralph H. Turner, Lewis Killian, and Neil Smelser to refer to social processes and events that do not conform to the established social structure but arise „spontaneously" and violates established norms in the social system. The research distinguished three main theories: (1) The Structural Approach, (2) Rational Actor and Resource Mobilization Theory, and (3) the Political Context and Political Opportunity Structures.

1.1.1 The Structural Approach

In 1959, Kornhauser published his *The Politics of Mass Society*, which emphasized the structural origins of revolutionary or protest movements and rapid social change as the catalyst for mobilization. Contributing to the high likelihood of individuals' behavior is the structure of society, which is characterized by the fragility of intermediary relationships, the compartmentalization of personal relationships, and the centralization of national relationships. Add to this the fact that the cultural characteristics of society tend to undermine multiple loyalties and that individuals lack a robust set of internalized norms, and it is easy to see why movements tend to lead to extremism and violence. The catalysts for these mass movements are rapid changes that reveal the vulnerability of political authorities, the absence of several independent groups, and the dismantling of previously established mediating relationships (Kornhauser, 2017).

In the theory of Smelser (1963), the trigger is caused by a breakdown such as an armed conflict or economic crisis. His value-added theory identified six essential components for social movements, which are structural conduciveness (things that make or allow certain behaviors possible), structural strain (injustice or inequality), generalized belief (explanation of the problem), precipitating factors, mobilization for action, and failure of social control (how the authorities react). He distinguishes, however, between collective outbursts of violence and collective movements, which direct collective efforts to modify norms. According to Smelser, social tensions must be combined with the appropriate structural conditions to lead to a social movement. In addition to these two components, there must be general beliefs triggered by the stresses. These elements combine to create a cause that the aggrieved mobilize to bring about change. In his theory, we could find examples of arrests or accidents that mobilized to action among the triggering factors.

1.1.2 Rational Actor Theory and Resource Mobilization Theory

Due to changes in the shape and content of social mobilization in the 1960s, the 1970s saw a deep review of interpretations of collective action. During this time, structural-functionalist theory lost favor, while rational election and organizational management models emerged.

These approaches were directly inspired by Mancur Olson's seminal work, *The Logic of Collective Action*. Olson's theoretical foundation is rational action theory or the economic model of man. He asserts that individuals who are either be "rational" or "self-interested" will not contribute to achieving their common goals. In this context, rational behavior implies that individuals' actions are guided by their costs and benefits and that individuals act in their own best interests. Individuals, in other words, tend to maximize their utility. It is critical to distinguish between utility maximization and self-interest. Individuals who are both self-interested and altruistic maximize their utility. The altruist finds fulfillment in serving others and improving their well-being (Olson, 1971). The intriguing statement in Olson's theory is that individuals are not envisioned as collaborating agents simply because they share shared problems or goals.

The Resource Mobilization Theory (RMT) has emerged as an alternative to Olson's model, rejecting both the irrational actor model and the deprivation and grievance caused by social strains (which was supported, for instance, by Smelser). According to resource mobilization theories, social movements' configuration depends on group resources such as human and material resources. In comparison with Olson's, Jenkins, McCarthy, and Zald, researchers at RMT, underlined that

1. actions are rational answers of adjustment to the costs and benefits of various lines of action;
2. changes in available resources, group organization, and opportunities for collective action; and
3. strategic factors and the political processes in which they are embedded.

As RMT researchers stated, mobilization is how a group gains collective control over the resources required to carry out collective action. It involves four main elements, foremost that (a) The organizational model is suitable for the situation, (b) It contains compensation for the costs of the mobilization (e.g., face-to-face interaction), (c) Recruiting focuses on solidarity incentives and capable of distributing the vision of change (d) The support of political elite or established organizations (Jenkins 1983; McCarthy & Zald, 1977).

The significant contribution of RMT was seemingly limited, but it caused a radical change from previous theoretical approaches because it set the action of social movements at the epicenter of research.

1.1.3 Political Opportunity Structures and its limitations

Kriesi introduced the Political Opportunity Structure (POS), claiming that the political system determines movement development, not the actors engaged (Benedicto & Kriesi, 1992). His idea was nicely incorporated by Tarrow into the analysis of social movements, claiming that the POS refers to the consistent dimensions of the political environment that encourage or discourage collective action, focusing on the external resources of the group. Social movements seize and expand political opportunities, transforming them into collective action and sustaining them through mobilizing and cultural frames (Tarrow, 1994). As a result of resource mobilization theory's emphasis on political context, we now better understand how collective action impacts social movements.

However, the most obvious limitation of the Political Opportunity Structures theory was that it overshadowed cultural and symbolic aspects and its production. Nor the POS nor the theses mentioned above could not provide an answer to the question of how do actors become part of collective action? How do they reach to recognize themselves in it? These queries were addressed, among others, by Melucci, who indicated the necessity of rethinking the concept of collective identity with a focus on the influence of culture (Melucci, 1980; Melucci, 1985; Melucci, 1988; Melucci 1989; Melucci 1995; Melucci, 1996).[1]

The analysis of collective action frameworks is one of the most widespread approaches that has moved social movement research away from the roots of Olson's rational choice and towards a broader collective logic of action (Snow et al., 1986; Snow & Benford, 1988; Hunt et al. 1994; Benford & Snow 2000). By maintaining the individual's emotional commitment to action, such framework work can help reduce resource costs. On the other hand, creating collective frameworks that are ideologically demanding, socially exclusionary, or conflictual fosters cracks and draws attention to how organizations manage or fail to overcome these differences. Moreover, resolving framework conflicts can involve mobilizing resources to bridge gaps between groups with different agendas and perspectives. Thus, while many strands of social movement theory have moved away from economic models of collective action, many still emphasize the need for organizations with close ties to their constituents and how identities are constructed and destroyed within such organizations and networks.

Formal organizations with resources are critical to recruiting and coordinating individuals for collective action. Researchers initially adopted the resource mobilization theory (RMT) and exemplified Olson's idea and his organization-centric solution. However, McCarthy and Zald (1973, 1977) emphasized external support and resources in the life of a social movement and focused on structural progress. Contemporary researchers went beyond Olson's oriented line of research and suggested that research

1 The works of Morris & Mueller (1992), Melucci (1995), and Baumgarten et al. (2014) also concentrated on how social movements embody themselves in specific cultural contexts and generate new cultural patterns outside of institutionalized environments.

should concentrate on the role of networks, culture, identity, political processes, and structures of opportunity, pointing overall toward a complex system (Melucci, 1996; McAdam et al., 2001; Della Porta & Diani, 2006).

2 Connective Action

Shortly after the first decade of the 21st century, the World started to learn of a new type of collective action mediated by modern technologies, social networks, and messaging applications. The digitally networked actions emerge as late modern democracies move away from parties, broad reform movements, and principles. In these circumstances, organizations have to discover that they need to involve people in other ways, treating them as partners rather than members and giving them unique opportunities to contribute and express themselves. However, this shift is mainly forced by the built-in features of social media. It also introduces micro-organizational resources in individual networks, content creation, and technology development skills. As a result, personalized collective action formations, in which digital media are integral organizational components, have joined, prevailed, and partially displaced collective action based on exclusive collective identifications and close networks. Of course, this also resulted in some of the resulting digital networks being adaptable to various conditions, problems, and scales.

As we have seen in the last part of this chapter, researchers early realized the new tendency, the new theory of this collective action, which adapted itself to this new era, was introduced by Bennett and Segerberg. They called it connective action as it reflects the increasingly complex and networked action in which numerous organizational structures utilize their resources off- and online, using various technologies in the distance (Bennett & Segerberg, 2012, Bennett & Segerberg, 2014). Moreover, connective action refers to an essential degree of technology-enabled networking created by digital media individuals to express their opinions directly to the world. However, the analysis of connective action requires focusing on different characteristics as it is, by its very nature, less similar to collective action.

One significant difference between collective and connective action is the structure. In Olson's thesis, collective action highlights the difficulties of getting individuals to contribute to the joint effort that typically involves pursuing some public good. In his idea, formal organizations with resources are crucial to recruiting and coordinating individuals in collective action. Connective networks, on the contrary, may show a very different structure, they can vary in terms of stability, coherence, and scale but are organized around different principles, and their actions tend to be more individualized. Moreover, web-organized strategies direct to action without demanding a collective identity or any organizational asset for answering an opportunity (Bennett & Segerberg, 2012, Bennett & Segerberg, 2014).

Collective action, moreover, demands various levels of mobilization of organizational resources employed in organizing, leadership, building common frameworks of action, and mediation to overpass differences in organizational structure. However, collective action today is typically characterized by a group of separate organizations that network to recruit and keep members and affiliated individuals in action. Therefore, personal interactions and connections may be seen as informal prerequisites for more centralized mobilization in collective action. The strategic task of mediating and building coalitions between organizations with different perspectives and interests moves into the focus of organizational analysis (Diani, 2011). Since the action dynamics of this logic do not alter essentially with social media, the focus should be on how these tools assist actors to do what they are currently doing (Bimber et al., 2009; Earl & Kimport, 2011, Bennett & Segerberg, 2012, Bennett & Segerberg, 2014).

These modifications in collective action are present, well visible in modern movements, and noticeable at an individual level. It is enough to consider how many protest events we could see monthly on our social media wall and how easily we could express our opinion towards these events in likes or shares (or ignoring). The classic form of mobilization has become more comfortable while seemingly do not require identifying attributes. People who were hard to reach up to this point and even harder to get to share personally transformative collective identities somehow managed to mobilize protest networks from cities to even countries. Indeed, social orientations are individualized, and people become structurally or psychologically unsuitable to modernist modes of political movement organizing. Organizing such individuals to overcome free riding and form common identities is not always the most successful or effective strategy. People who seek more individualized paths to concerted action are already familiar with a different organizational form: connective action.

From sociology's point of view, the connective action underscores a new dynamic, recognizing digital media as organizing agents. Researchers have examined how digital communications technology changes Olson's basic idea of collective action. Weak individual engagement in large organizations (free riders) may not be as accurate as Olson suggested (Lupia & Sin, 2003). Former free riders may find it easier to join political networks that blur the lines between public and private, according to Bimber et al. (2005). Benkler's explanation of digitally mediated social networks may be decisive for this objective (Benkler, 2006). Participants become self-motivated when their content is shared and recognized by others, who then repeat the connected sharing behavior. Individuals can cooperate when empowered by technology platforms that coordinate and scale interpersonal networks without formal organizations or the shifting of social identities. Social networks involve co-production and co-distribution, revealing a new economic and psychological logic: co-production and sharing based on a personalized expression. This does not apply to all online conversations. The logic of the organization-centric stationary world is often mirrored online, with the organizational logic changing little beyond possible efficiency gains (Bimber & Davis, 2003; Foot & Schneider, 2006).[2]

In the connective action, sharing ideas and behaviors in trusting relationships becomes an act of personal expression and recognition or self-affirmation, even if the people involved are on the other side. They can connect without any shared framework, such as club membership or ideology. A self-motivated sharing of internalized or personalized ideas, plans, images, and resources with networks of people replaces the initial difficulty of collective action in motivating individuals to contribute. This sharing can be done on social networking sites like Facebook, or more public platforms like Twitter and YouTube, through comments and retweets. The effect of action networks of this logic can work as snowball thanks to quickly disseminated personal action frameworks and digital communication technology (Bennett & Segerberg, 2012).

The network functions as an organizational structure in this new form of collective action. Indeed, technology-enabled communication networks involve more than just the transmission of messages. However, this online network and its offline extensions are more than just communication systems. These networks are fluid groups that can quickly adapt to changing political goals, often crossing geographic and temporal boundaries and

2 It also helps explain why people collaborate for free to build things like open-source software (Calderaro 2011).

distributing news and pieces of information even faster. According to Diani (2011), networks are organizational systems that can go beyond the essential components of organizations and individuals; while communication technologies do not change the collective action logic of large networks, they do this by connecting networks of action.

Using Latour's (2005) actor-network theory, we can identify digital network systems (e.g., various social media and devices that power them) as potential network agents alongside human actors (i.e., individuals and organizations). Organizational connectors (web links), event coordination and information exchange sites (Youtube, Facebook, Twitter), and multifunctional network platforms (links in Twitter and Facebook postings) are examples of digital mechanisms. These technologies help calibrate connections by creating openness, privacy, security, and interpersonal trust. Furthermore, these digital traces must remain on the web to create memory records or repertoires of actions that can be passed on through traditional collaborative action methods, such as rituals or formal documentation.

Since collective and connection action are different action logics concerning the consideration of identity and electoral processes, they require a separate examination. As with conventional collective action initiatives, the results of digitally mediated networking processes are not predetermined. Based on the types of social technologies being built and appropriated by participants and the possibilities that can evoke anger or compassion in large numbers of individuals, the transmission of personal expression over networks can become scalable, stable, or robust. Occupy Wall Street might not have grown so large without inspiration from the Arab Spring, the outrage in Spain, or the deteriorating economic conditions that fueled the anger of displaced people around the world. Occupy movements, for instance, were centered around a few recognized mainstream political organizations. Leaders and official spokespeople were avoided at all costs. The most visible forms of organization were layers of social technology and websites that contained attendee messages and customized networking tools.

Furthermore, as Bennett and Segerberg (2012) remarked, collective and connective action can coincide in the same action ecosystem. The different types of connective action networks, unifying networks of action, are characterized by the involvement of formal organizations in supporting individual engagement. As we have seen, traditional organizations are less critical because social technologies may work as self-organizing networks, such as the Indignados movement in Spain, the Arab Spring revolutions, or global occupation rallies. On the other hand, hybrid networks may exist where traditional organizations operate behind the protest and set up advocacy networks to enable individualized interaction. These characteristics indicate a three-part typology in which coexistence, stratification, and transmission or evolution may also exist. However, this could be a part of a natural evolution of a movement that derives from social dynamics.

Table 1: Comparison of the Collective- and Connective action

	Collective action	**Connective action**
Organization	Solid organizational coordination	Individual coordination, where the social network is the organizing agent
Structure	Solid link between actors	Weak formal connection between the actors
Motivation	Creating a common purpose to achieve a collective benefit as a group	Engagement is self-motivated where the sharing of information along with acknowledging and circulating the shared material is fundamental
Identity framing	The identity framing is based on a particular set of beliefs and symbols („we are")	The identity framing is not based on a particular set of beliefs and symbols
Communicative content	The structure of collective action is based on the framing of collective identity, the communication revolving around it	In the context of connective action, personalized action frames provide easily transferable communication that is a subject of free interpretation, including asynchronous non-verbal communication (e.g., memes, likes, sharing)
Limits	Limited by a defined obligation or by a restrictive policy	Unlimited connectivity where extensive and flexible networks replacegroup relationships.

2.1 A typology of collective and connective action networks

The logic of connective action applies these different action logics to construct a three-part typology for the dominant protest networks in recent politics: Crowd-Enabled Connective Action, Organizationally Enabled (Hybrid) Connective Action, and Organizationally Brokered Collective Action. Although all three models rely on and are framed by digital technologies assisting people to personalize their political participation, the models also help to describe differences and dynamics within large-scale action networks in event-centric disputes, such as protests and protest sequences. The models can also apply to more robust advocacy networks, such as campaigns, that engage individuals in supporting causes daily. However, it should be noted that this is not an attempt to comprehend, much less resolve, the many disputes among social movement scholars.

2.1.1 Crowd-Enabled Connective Action

A crowd-enabled connective action is a well-known form of collective action. The main characteristic is that crowd-enabled connective actions are organized without central or leading organizational actors and are less consciously created. However, in this case, the individual action frames replace collaborative action frames as transmission units over trusted social networks, while informally organized groups are more peripheral or exist both online and offline. In the approach of Mueller (2010), this is called an „associative cluster," thus emphasizing the relative lack of existence of purposeful organization and actors. However, as Bennett and Segerberg noted, even though some groups may have small media resources and limited organizational characteristics, these could mark a growing incidence of a connective organization (Bennett & Segerberg, 2012, Bennett & Segerberg, 2014).

2.1.2 Organizationally Enabled (Hybrid) Connective Action

The hybrid term was first used by Chadwick, who found that many organizations (e.g., non-governmental organizations, interest advocacy groups) do not necessarily fit into fixed categories (Chadwick, 2007; Chadwick, 2013). Instead, they mostly range from issue advocacy groups to policy think tanks, SMOs that run campaigns or protests, multi-issue groups, and places where people can connect and act together. One of their characteristics is changing their organizational repertoires to move from being hierarchical, mission-driven NGOs to being facilitators in loosely linked public engagement networks to get big groups of people involved (Bennett & Segerberg, 2012, Bennett & Segerberg, 2014).

Formal organizational actors use resources to implement social technologies that enable loose public networks to grow around individualized action themes. The organizationally enabled (hybrid) connective action may include more informal organizational players without collective identities. It is enough to have consistent attendance, donation, food distribution, or action coordination. Moreover, the more extensive communication networks surrounding these nodes considerably increased the influence of the network. It was noticed that many of the most devoted protestors spent long hours with declining numbers of peers discussing ways to boost participation without compromising the levels of dedication and activity that they believed vital to their value system. The hybrid-type action engages people on causes that may not be as appealing when a more prominent member or collective demands go hand in hand with the organizational offerings. These networks often use a mix of purpose-built and outsourced communications technology.

Bimber et al. (2005) observed advocacy organizations needing to develop looser, more entrepreneurial relationships with supporters; the hybrid-type formations reflect the pressures. Campaigns, protests, and advocacy networks on everyday issues, all sharing common organizational signatures such as

1. trusted NGOs and other civil society organizations loosely banding together to provide a network backbone,
2. digital media networks that engage the public in contentious political issues, and
3. minimal effort to brand the problems of particular organizations.

In addition to promoting their political agendas, the organizations acted as members of issue networks and facilitated personal engagement through easy-to-share images and individual frames of action.

Personalized connective action networks intersect with conventional collective action networks focused on SMOs, advocacy groups, and NGOs. As a result, while an organizational form in themselves, these networks are often challenging to capture and analyze. In most cases, formal organizations are centered, hierarchical, based on mission and territory, and defined by known and countable members. As a result of multiple centers, many of today's issue and cause networks are decentered, distributed, or organizationally flattened, geographically, and issue-bounded. Dynamic about changing populations that may turn on or off as different engagement opportunities are presented (Bennett, 2003; Bennett, 2005).

2.1.3 Organizationally Brokered Collective Action

The main characteristic of the organizationally brokered collective action is a large-scale action network that relies on intermediary brokering organizations, facilitating collaboration and bridging differences. These networks may use social media mainly for mobilizing and organizing participation and coordinating goals. The conventional sociological research about social movements, such as McAdam (1986), Keck and Sikkink (1998), created descriptions of NGOs that fit this category (Bennett 2005). We find the Stop Climate Chaos, Put People First network or the G20 Meltdown among the organizationally brokered collective action group.

Overall, Bennett and Segerberg’s three-part typology approach is not without limitations. There is a possibility that others will consider such networks as noise. However, careful analysis can help understand the democratic potential and effectiveness of connective and collective action instances. Digitally networked action is often more than clicktivism or the cumbersome organizational outsourcing of social networking to various commercial sites. The basis of the unifying action is sharing: the individualization that drives action and content on social networks.

In addition, modern communication technologies, social networks, chat applications enable the construction and stabilization of network structures. The technological agents that enable sharing in these contexts crowd out the free riders. In addition, the connective action focuses on the action dynamics of recombinant networks so that networks and communication are more than just prerequisites and information. This organizing principle differs from notions of collective action, which are based on core assumptions about resource roles, network roles, and collective identity. The great challenge of this typology is the distinction between unorganized and productive digital networked action. As has been shown, the political capacities of networks differ depending on whether or not the network has a stable core of organizations that share communication links and employ high levels of personal engagement mechanisms. When the digital networks are redundant and dense with paths for individuals, the networks have converged and allow the transmission of personal intervention mechanisms. Connective action does not explain all political controversies and does not replace the classic Collective Action model for studying social movements. Nonetheless, it illuminates a crucial protest course in today's polarized society and politics.

3 Related work about social media

The primary purpose of the following part of the literature review is to classify the published protest-related case studies according to the theory of collective and connective action and reveal the main research tendencies and gaps in contemporary scholarship. The subsequent examination encompasses a broad spectrum of protests and activism, including anti-austerity movements, Green politics-based protests, or civil disobedience. Chronologically we will discuss the followings: Moldova civil unrest (2009), Student protests in Austria 2009, Israel-Gaza (2009), Iran election protest (2009-2010), Tunisian revolution (2010-2011), Venezuelan protests (since 2010), Protest against the Stuttgart 21 development project (2010), Toronto G20 (2009), Tahrir Egypt (2011), England (2011), Occupy Wall Street (2011), Indignados (2011), Aganaktismenoi, Anti-austerity movement movements, Greece (2011), Italy (2011), Wisconsin labor protest (2011), Israel Hamas demonstration (2012), Brazil Vinegar (2012), Protest against the Taksim Gezi Park development project (OccupyGezi), Turkey (2013), Euromaidan (2013-2014), Black Lives Matter (2013), Romanian protests (2017–2019), Serbian protests (2018–2020), Hong Kong protests (2019–20).

3.1 Moldovan parliamentary election protests and civil unrest (2009)

The indicator of the civil unrest was the Moldovan parliamentary elections in 2009. Prior to the official results being announced, protests began on April 6, 2009, in major cities, including Chișinău and Bălți, Moldova's second-largest city, and lasted almost a week (6–12 April 2009). The demonstrators demanded a recount, a new election, or the government's resignation, claiming that the elections were fraudulent. In the initial phase of the protest, NGOs (ThinkMoldova, HydePark) collected contacts of potential protesters and created social media discussion groups where they could communicate. Students and NGOs who joined forces to create a new group used Twitter to organize a flashmob protest in Chișinău on April 6, a more significant protest and riot on the next day. Over 30,000 protesters, primarily students and young people, turned violent as police used tear gas and water cannons. They attacked the parliament and the president's office and set fire to documents and furniture. Police retook the building later in the evening. In the classification of connective action, the Moldovan civil unrest was a hybrid, organizationally enabled connective action as it evolved from a flashmob organized by Twitter, social network while NGOs provided know-how and organizational background (van der Zee, 2009, Lysenko & Desouza, 2012).

During the protest, Lysenko and Desouza (2012) analyzed the role and use-cases of the internet– and cellular-based communication technologies. They find that protesters organized their initial mobilization through Twitter and short message service (SMS). However, as they examined, Twitter was mainly used during the later stages of the uprising and the subsequent information war for communication about the local and global conflict. According to Lysenko and Desouza, through skillful use of new Internet-based communication technologies, it is possible to conduct a successful revolution without noticeable prior offline organization. The limitation of their studies is that they manually analyzed a limited number of tweets; however, they established the administrative role of NGOs in the protest.

3.2 Student (Uni brennt) protests in Austria in 2009

Students occupied several the lecture halls of several Austrian universities to protest against tuition fees on 20 October 2009. The protests were the largest in recent years in Austria and sparked a broad debate about education policy. Overall, 11 university locations (Vienna, Graz, Linz, Klagenfurt, Innsbruck) across Austria were occupied by students. Education, politics, civil society, trade unions, the arts and culture, and the media all expressed some solidarity. The protesters are democratically organized and used social media such as Twitter and Facebook to communicate, so their communication was based on the individual's interconnected individual and personal networks. Aside from the demonstrations, various working groups were formed. Plenaries and cultural and educational events were held in the occupied rooms. The protesters want universities funded and democratically run and tuition fees abolished or not introduced.

For the protest, Maireder and Schwarzenegger conducted a qualitative analysis of the communication and examined how those involved used social media. Citing the concept of voluntary issue communities, they argued that the low level of commitment and commitment, as well as the immediate affiliation with the movement, was fundamental to its success in reaching large groups of students and university staff, despite the protests and the broad public attention that the movement finally received. According to the authors, this movement of networked individuals is an example of how individuals can organize themselves as a politically active community. In addition, they pointed out what such a community and its actions can look like when participants naturally use information and communication technologies to connect socially within networked publics. Manually selected tweets and other messages served as data for their analysis.

3.3 Iranian presidential election protests 2009-2010

The so-called Green Revolution broke out in June 2009 to protest against the results of the Iranian presidential election, which President Mahmoud Ahmadinejad won. The demonstration spread fast to Iran's major cities and lasted 7 December 2010. The government used the police and paramilitary Basij to suppress the protests; protesters were beaten, pepper-sprayed, arrested, tortured, and even shot in some cases. The universities of Tehran were closed, websites were blocked, and rallies were banned. Although the Iranian Green Movement arose during these protests, the protests were quelled.

Wojcieszak and Smith (2014) analyzed the events in the scientific literature. They surveyed 2800 young, educated, metropolitan, and technologically savvy Iranians over a year after the election to examine what sources these youth use for information, the extent to which they rely on new media for political exchanges, their experiences with online censorship, and political efficacy as related to new media. Although the authors are skeptical about technology's ability to sustain revolution, they also identified what can be described as hubs of politicized Iranian youth. Data for this analysis come from an online survey targeting Iranian youth.

3.4 Tunisian revolution (2010-2011)

The Jasmine- or Tunisian Revolution broke out on 17 December 2010 when Mohamed Bouazizi, a local vendor, committed self-immolation. The protests, which lasted 28 days, were Tunisia's most violent social and political, with scores of deaths and injuries caused mainly by police and security forces. They forced Ben Ali to resign on 14 January 2011, after fleeing to Saudi Arabia, ending his 23 years in power. The protests eventually led to a complete democratic transition and free elections and aided in building a pluralistic democracy in Tunisia. The Tunisian success sparked a chain reaction known as the Arab Spring movement.

Lotan et al. (2011) were among the first who tried to analyze the dissemination of news on Twitter during the events. The researchers' approach aimed to differentiate the users (bloggers, activists, journalists, and mainstream media outlets) and analyzed patterns of sourcing and routing information among them. Using this analysis, the Lotan et al. discussed how social media played a crucial role in intensifying and distributing timely information globally. However, the notable limitation of their studies was that the analysis focused solely on English messages and news.

Just a few years ago, Kharroub and Bas (2016) recognized that the role of images in social media was under-researched. Their study analyzed the potential of emotional and impactful visual content to motivate activism, focusing on the 2011 Egyptian revolution. Without normalizing the data, Kharroub and Bas emphasized that images posted by users outside of Egypt received the most attention, while the weight of attention might correlate with the number of followers of the user who posted the particular image. That limitation aside, their study emphasized the impact of visual content on social media audiences.

3.5 Socioeconomic crisis in Venezuela (since 2010)

The Venezuelan crisis is an ongoing socio-economic and political crisis that began under Hugo Chvez and worsened under Nicolas Maduro. Hyperinflation marked it, increasing hunger, disease, crime, mortality, and mass emigration. Venezuela's shortages prompted Chvez to declare economic war on June 2, 2010. The crisis worsened under the Maduro government due to low oil prices in early 2015 and a lack of maintenance and investment. Faced with falling oil revenues, the government has responded by denying the crisis and violently repressing the opposition. Additionally, extrajudicial killings by the Venezuelan government have become common, with the UN reporting 5,287 by Special Action Forces in 2017 and at least 1,569 in the first half of 2019 (Webber, 2010, Larmer, 2018).

Morales, Losada, and Benito (2012) analyzed Twitter data to describe the structure and dynamics of the emergent social networks during the political crisis. They focused on a

Venezuelan protest by Twitter in December 2010 and discovered that a community structure with highly connected hubs and three different kinds of user behavior define the information flow dynamics without examining the message content. The data was received by Twitter API and consisted of 421,602 messages that included the hashtag #SOSInternetVE using the Twitter API interface between 14–19 December 2010.

The COVID-19 epidemic made a turning point in the socio-economic crisis in Venezuela. Munger et al. (2019) applied a topic modeling on the tweets of Venezuelan politicians during the 2014 anti-Maduro rallies in Venezuela and studied trends in hashtag use by the opposing coalitions. The researchers suggested that the regime's best approach to existential danger is to advance numerous rival narratives addressing unrelated topics to the opposition's criticism. Furthermore, their findings indicate that the two coalitions used distinct rhetorical methods, which corresponds to our predictions about how to manage the protesters' disagreement.

3.6 Protest against Stuttgart 21 railway project (2010)

The Stuttgart 21 protest was against the Deutsche Bahn project to rebuild the Stuttgart railway junction. Subway stations were built to connect the new Stuttgart – Wendlingen line to the Stuttgart central station. The criticisms included a lack of democratic legitimacy and citizen participation, security flaws and traveler access issues, high costs and inefficiency, a threat to mineral water resources, the inefficiency of the new central station, and planning flaws. Other ideas were discussed as terminus 21 and transfer 21. A non-violent protest is expressed through petitions, information stands, demonstrations (especially the weekly Monday demos). Protests against the project started in 1996, where participants range in age, education, and profession. The action alliance against Stuttgart 21 and the park guards were the most well-known S-21 opponents. On 30 September 2010, a police operation in the palace gardens drew national attention to the protest movement. The project's proponents and opponents held live online and television arbitration talks. After the 2011 state elections, protests against Stuttgart 21 influenced the first green mayor of state capital election in October 2012. The green-red state (Social Democrat/Green) government mediated the conflict on 27 November 2011, bypassing the Stuttgart 21 referendum (The Gurardian, 2010, RND/DPA, 2020).

Jungherr and Jürgens (2014) used Twitter to analyze the protests against the infrastructure Projekt Stuttgart 21. They focused on the hashtag #s21, posted between May 25, 2010, and November 14, 2010, by the 80,000 most followed users in Germany. The dataset consisted of 165,059 messages. They distinguished events that resulted in high activity and the changes in user behavior patterns that varied from their usual patterns and revealed those artifacts (links) that dominated conversations during times of high activity indicative of tactical support of the protests.

3.7 G20 Toronto summit protests (2009)

Protests and demonstrations began one week before the 2010 G20 Toronto summit on 26–27 June in Toronto, Ontario, Canada. Protests included anti-poverty and anti-capitalism. They were mostly peaceful, but a group of protesters using black bloc tactics vandalized several businesses in Downtown Toronto. More than 20,000 police, military, and security personnel patrolled the 10,000-strong protests. At least 40 shops were vandalized, causing at least C$750,000 in damage. Over 1000 people were arrested, making it Canada's largest mass arrest. Police brutality during arrests was heavily criticized after the protests by the media and human rights activists. Despite the Toronto Police's repeated attempts to stop court proceedings by appealing the case, a class-action lawsuit was filed on behalf of all those arrested (Douai, 2014, The Canadian Press, 2020.).

Earl, Hurwitz, Mesinas, et al. (2013) used Twitter to analyze protests surrounding the G20 meetings held in Pittsburgh, PA, in September 2009. Examining the content of 30,296 tweets over nine days, they developed hypotheses about the content of tweets during protests, arguing that Twitter was used to reduce information asymmetries between protesters and police. They revealed that protesters frequently used Twitter to communicate information, including details about protest locations and the sites and actions of police, which is information that was formerly monopolized by the police, thus creating a new dynamic in protesters and police interaction toward information symmetries. Earl, Hurwitz, and Mesinas' study is based on manually coded tweets.

Using Twitter as an example, Poell and Borra (2012) examined the appropriation of social media as platforms for alternative journalism by protesters at the 2010 G20 summit in Toronto, Canada, coordinated by the Toronto Community Mobilization Network. The

authors analyzed 11,556 tweets, 222 videos, and 3,338 photos with the hashtag #g20report on Twitter, YouTube, and Flickr because the network encouraged participants to share messages through these social media sites. Data was collected for the 12 days between 22 June and 3 July. According to the authors, the results suggest that social media has not facilitated the crowdsourcing of alternative reporting, except to some extent for Twitter.
The researchers classified their data manually using emergent coding. It was first examined, and a checklist of distinguishing characteristics comprised of search terms and photos was compiled. Then, the associated characteristics were classified (e.g., police activity, protestors' issues, black bloc, condemning violence, and arrests). Following that, the material was coded according to these categories.

In another paper, Poell (2014) used the same dataset and focused on how the massive use of social media in recent protests affects the character of activist communication? Through his analysis, he traced the hyperlink network that included protest communication. The hyperlink analysis sheds light on the internet environment where this communication took place. Additionally, the investigation examines how the various social platforms' technology designs, associated user practices, and commercial models shaped communication. This analysis demonstrated that social media usage accelerates activist communication and significantly enhances its visual quality. Additionally, as activists embrace corporate social media on a large scale, they lose control over the data they generate collectively and the basic designs of the places they communicate.

3.8 Egyptian revolution (2011)

Tahrir Square was the epicenter of the 2011 Egyptian revolution against Mubarak, which started on 25 January, when over 50,000 protesters occupied the square. In the days that followed, Tahrir Square remained the focal point of protests in Cairo. On 29 January, Egyptian fighter jets flew low over the square. At least 100,000 protesters were reported on 30 January, and at least 250,000 on 31 January by Al Jazeera correspondents. On 1 February, Al Jazeera reported that over a million peaceful protesters gathered in the square and nearby streets. However, according to Stratfor's analysis, the actual number of protesters never exceeded 300,000. 9 February 2011 – Tahrir Square The square became a symbol of the ongoing Egyptian democracy demonstrations (Stratfor, 2011). On 2 February, pro-Mubarak and pro-democracy demonstrators clashed in the square, followed on 3 February by the ‘Friday of Departure’ demonstration, one of the named “day of” events. Tahrir Square’s image and the name became global within a week of international media coverage. During the uprising, a staff of twenty people ran a Facebook page called “Tahrir Square” to counter the lack of or distorted coverage of events and responses in state-run and state-aligned media. The 18-day revolt in the square allowed the Egyptian Armed Forces to depose Mubarak on 11 February 2011 (Vatikiotis, 1997; Hamanaka, 2020).

Aday and colleagues (2013) examined consumption trends of Arab Spring-related content using a unique data set created by integrating archived Twitter content with metadata extracted from the URL shortening service Bit.ly. The researchers examined tweets with the hashtags #sidibouzid (Tunisia; 79,166 total tweets), #jan25 (Egypt; 665,092 total tweets), #feb14 (Bahrain; 48,015 total tweets), and #feb17 (Libya; 885,724 total tweets). They tried to answer who visited the links posted on Twitter? Additionally, who received the most online attention: demonstrators and other non-elite individuals or established news organizations? Their data indicated that the great majority of attention paid to Arab Spring content originated outside the MENA region and that mainstream media, rather than citizen media, dominated the global conversation throughout the uprisings. Consequently, the researchers concluded that Twitter was generally beneficial as an information route for non-MENA observers but less so for on-the-ground protesters.

Burns, Highfield, and Burgess (2013) conducted a study to determine how these Arab and English groups interacted directly (especially taking into account possible language barriers between them). Using hashtag data collected between January and November 2011, her article examines Twitter usage patterns between the popular revolution in Egypt and the civil war in Libya. They examine the volume of tweets sent by English, Arabic, and mixed-language Twitter users over time and the networks of contacts between these groups as they formed and changed during these revolutions, using tools specifically designed for large amounts of data used. They discovered typical patterns of information flow between the English- and Arabic-speaking portions of the Twittersphere by examining reply and retweet traffic and highlighting the user's responsibility in bridging the two language domains. Researchers used the Twitter Application Programming

Interface (API) to track #egypt and #libya from early 2011 (January 23, 2011, for #egypt; February 16, 2011, for #libya); our data collection period for their analysis ended on November 30, 2011.

Choudhary, Hendrix, Lee et al. (2012) analyzed over 800,000 tweets on revolution-related topics to present their findings in three different ways: First, they examined how sentiment evolved in response to evolving events; how the most influential tweeters and most popular tweets shed light on the most influential Twitter users and the types of tweets that resonated the most; and how user sentiment and follower relationships relate to dynamic properties and sentiment of social networks. Then, they compared Egypt-related issues during the revolution to other prominent trending issues in early 2011 using these measures (such as politics, sports, and entertainment). The authors noted that the discussion was less coherent than other Twitter topics due to the intense negative sentiment. Additionally, a significant portion of the debate reflected the broadcast news coverage of ongoing events, with the most influential users and tweets delivering news and a significant portion of the news urging people inside and outside Egypt to re-tweet the news.

Doerr, Fouz, and Friedrich (2012), concerning the events of Arab Spring, tried to find the answer to why rumors spread quickly on social networks. The researchers simulated a natural process of rumor propagation on graphs representing real-world social networks and several classic network topologies. They also performed a mathematical analysis of the process. Both simulation and analysis demonstrate how rumors spread on social networks. An important observation in the mathematical proof and a reasonable explanation for this phenomenon is that small-degree nodes learn a rumor as soon as one of their neighbors knows about it and quickly propagate it to their neighbors. This propagation scheme facilitates the sending of rumors from one high-degree node to another. According to the authors, it partly explains why social networks disseminate information quickly, even though the process is not centrally organized and not intelligently designed—fruitful interaction between hubs with many connections and average users with few friends matters. Hubs make the news available to a large audience, while average users quickly transfer the information from one neighbor to the next. For example, according to the authors, a rumor started at a random Twitter node, reaching 45.6

million of 51.2 million users in only eight rounds. Overall, Doerr and colleagues further assisted the understanding of the effects of rapid news distribution.

3.9 England riots (2011)

The 2011 London riots were a series of riots between 6 August and 11 when thousands of people rioted across England, resulting in looting, arson, police brutality, and five deaths. Protests began in Tottenham Hale, London, after a local man, Mark Duggan, was killed by police on 4 August. After Duggan's death, several violent clashes with police resulted in the destruction of police vehicles, a double-decker bus, and many homes and businesses. Overnight, looting occurred in Tottenham Hale and Wood Green. The following days saw rioting in Hackney, Brixton, Walthamstow, Peckham, Enfield, Battersea, Croydon, Ealing, Barking, Woolwich, Lewisham, and East Ham. A spate of "copycat violence" occurred across England from 8 August to 11, with social media playing a part in some cases. By 10 August, over 3003 people had been arrested across England, with over 1984 facing charges related to the riots. Local economic activity, which was already struggling due to the recession, was severely harmed, causing an estimated £200 million in property damage. The riots sparked intense political, social, and academic debate about their origins and context. The rioters' actions were blamed on racial and class tensions, economic decline, and the resulting unemployment (BBC, 2011, Mirror, 2011, Moran & Waddington, 2016).

Tonkin, Pfeiffer, and Tourte (2012) investigated 600,000 tweets and retweets about the riots to determine whether Twitter was a primary organizing tool for illegal group action. The findings suggested that unimportant tweets faded and that Twitter users retweeted to reinforce their beliefs in the context of others' comments. In addition, tweets from well-known individuals received a higher rate of retweets. In the case of the British riots, there is less evidence that Twitter was used to incite criminal activity at the time; however, it did aid in disseminating information about the following events.

Procter, Vis, and Voss (2013) analyzed crisis communications during the protest and confirmed that police, emergency services, and government agencies faced complex problems using social media platforms effectively, as rumors on Twitter were self-correcting. The authors emphasized a need to provide information and advice more timely, from sources that the public can trust. Their results indirectly also showed how effective the misinformation on social media could be.

Panagiotopoulos, Bigdeli, and Sams (2014) studied the dynamic elements of collaboration during the 2011 English riots using a dataset of 1746 tweets from 81 local government Twitter accounts. They searched for answers to how local governments tried to reduce rioting and help community healing. The authors created a range of valuable and actionable messages using Twitter's conversational and fast updating features. In several cases, the collective against riots evolved mutually: local officials mobilized residents, and citizens mobilized local authorities.

3.10 Occupy Wall Street, anti-austerity movement (2011)

Occupy Wall Street began in September 2011 in Zuccotti Park, New York City's Wall Street financial district, against economic inequality and money's influence in politics. It sparked the Occupy movement in the US and abroad. The protest was called by the anti-consumerist magazine Adbusters. One of Occupy Wall Street's main concerns was the undue influence of corporations on government, particularly in the financial services sector. It refers to the income and wealth disparity in the United States between the wealthiest 1% and everyone else. Protesters achieved their goals by following consensus-based decisions made in general assemblies that favored direct action over petitioning authorities. They were evicted from Zuccotti Park on 15 November 2011. Protesters then occupied banks, corporate headquarters, board meetings, foreclosed homes, and college campuses (OccupyWallSt, 2011, Berrett, 2011).

In their work, Bastos, Mercea, and Charpentier (2015) identified around 100 Twitter hashtags related to the Indignados, Occupy, and Vinegar protests and over 100 Facebook sites and groups dedicated to the events. Using time-series data from Twitter, Facebook, and onsite demonstrations, they evaluated the Granger causality between social media streams and onsite events during the Indignados, Occupy, and Brazilian Vinegar protests. They discovered that heated discussion on Twitter and Facebook predicted onsite protests during the Indignados and Occupy movements, with bidirectional Granger causation between online and onsite protests in the Occupy series. On the other hand, the Vinegar demonstrations established Granger causality between Facebook and Twitter communication and between demonstrators and onsite injuries and arrests. They concluded that anticipating protest action is likely to vary among examples of unrest.

Gleason (2013) studied how people learn about Occupy Wall Street via the microblog Twitter. He investigated the existence of informal learning about Occupy Wall Street using descriptive statistics, content analysis, and a case study. According to Gleason, Twitter offers numerous ways to participate in the Occupy movement, from creating, tagging, and sharing information to reading, watching, and following a hashtag. As he stated, Twitter's social and technical characteristics are offered to assist ground this debate.

Park, Lim, and Park (2015) investigated the network structure, interaction pattern, and geographic distribution of 328 Occupy Wall Street activists on Twitter and YouTube. They constructed a loose hub-and-spoke network, implying that a few central users managed information and bridged tiny communities. YouTube videos with similar themes produced a dense mesh network, reinforcing common meanings. By geographic distribution, both Twitter and YouTube networks were actively organized by Americans, although anonymous individuals primarily engaged on YouTube. These findings demonstrate how Twitter not only organizes and coordinates information but also fosters the flow of ideas between groups. Thus, YouTube is perfect for spreading ideas and strengthening member solidarity. The findings have implications for the various functions of social media platforms in mobilizing collective action.

A typology of the use of Twitter in this protest was analyzed by Penney and Dadas (2014), which highlights its use in connection with direct action. Their study used 17 in-depth interviews with Occupy Wall Street activists (10 men and seven women) from different geographic regions (including New York City, Boston, Chicago, Austin, Portland, Seattle, Helena, and other parts of California). Her example shows how protesters can quickly build a geographically dispersed, connected counter-public that can communicate critiques of authority outside the confines of mainstream media. They also discovered that sharing pre-existing content was seen as making as much sense as composing unique works. As a result, these Twitter users increased the transmission of information and developed and supported an OWS counter-public. However, relying on an external platform exposes protesters to communication restrictions and unwanted surveillance by unfriendly authorities.

Tremayne (2014) used network analysis to answer the following questions: What were the main OWS Twitter hotspots in 2011? How did OWS arise from a slew of social movements spearheading a nationwide protest? What critical moments in the Twitter conversation aided the scale shift? His study combines social movement notions with network centrality measurements by addressing these problems but bearing considerably spatial limitations.

3.11 15-M, Indignados, Anti-austerity movement (2011)

The 15-M Movement and the Indignados Movement were anti-austerity protests, demonstrations, and occupations in Spain that began around the 2011 and 2012 local and regional elections. Starting on May 15, 2011, many subsequent demonstrations spread through social networks like Real Democracy NOW and Youth Without a Future (Spanish: Juventud Sin Futuro). The movement was linked to the Spanish financial crisis of 2008–14, the Arab Spring, and protests in North Africa, Iran, Greece, Portugal, and Iceland. The movement was also compared to Stéphane Hessel's political manifesto Time for Outrage!, which aimed to empower unemployed or underemployed Spanish youth. Protesters in Spain chanted against high unemployment, welfare cuts, politicians, the two-party system, capitalism, banks, and public corruption. Many demanded fundamental rights to housing, employment, culture, health, and education. The movement brought the Arab Spring protest camp model to Europe, adapting it to a more countercultural framework. This would later grow to influence Occupy Wall Street. According to RTVE, between 6.5 and 8 million Spaniards participated in these events (Castañeda, 2012, Nez, 2021).

González-Bailón, Borge-Holthoefer, and Moreno (2013) examined the growth of online mobilizations using data from the Indignados (outraged) movement in Spain, which arose under the influence of the revolution in Egypt and as a precursor to the global Occupy mobilizations. Their data tracked Twitter activity surrounding the May 2011 protests, which led to the formation of campgrounds in dozens of cities across the country and massive daily demonstrations in the week leading up to the 22 May election. They reconstructed the network of tens of thousands of users and monitored their news activity for a month (25 April 2011 to 25 May 2011). Based on the structure of the network and the level of activity in messaging, they identified four types of users and analyzed their role in the growth of the protest. Starting from theories of online activism and research on information dissemination in networks, with the following two questions: How does protest information spread in online networks? Moreover, how do different actors contribute to the growth of activity? These questions informed the theoretical debate on whether digital technologies are changing the logic of collective action and provide evidence of how new media facilitate the emergence of massive offline mobilizations.

By examining the usage of Twitter by the 15M movement, Peña-López et al. (2014) studied the nature of networked citizen politics as a different representational sort of political engagement. They started it by defining users, including how motions spread amongst them. Then they looked at the associations between networked citizen movements and formal democratic institutions, particularly political parties and mass media movements. They also looked at how networked citizen politics may use techniques comparable to Politics 2.0 but with different goals. Finally, their research demonstrates that individuals, behaviors, and ideas are exchanged and used as building blocks for further activity despite the lack of traditional organizations. Exceptions are minor and left-wing parties, but there is little inter-institutional dialogue.

3.12 Aganaktismenoi, Anti-austerity movement, Greece (2011)

Anti-austerity protests and general strikes took place across Greece. The protests began on 5 May 2010 in response to plans to cut public spending and raise taxes for a €110 billion bailout to resolve Greece's debt crisis. One of Greece's largest protests since 1973 claimed three lives on the protest of 5 May. On 25 May, the protesters began demonstrating in major Greek cities. Unlike previous protests, this second wave began peacefully and was not partisan.

Nevertheless, some of the events turned violent, particularly in Athens. On 7 August, demonstrators were removed from Thessaloniki's White Tower Square by municipal police. On 29 June, riot police and activists clashed as the Greek parliament voted to accept EU austerity demands. International media outlets (e.g., BBC, CNN, The New York Times) and academic research and Amnesty International reported on police brutality. The anti-austerity movement of Spain partly inspired the protests in Greece (Dalakoglou, 2013, Sotiropoulos, 2017).

Theocharis, Lowe, van Deth, and García-Albacete (2015) asked to what extent the anti-austerity movement movements used social media, among other things? To answer this question, they conducted a comparative content analysis of tweets sent during the heyday of each campaign. Although Twitter was heavily used for political discussions and to transmit protest information, the results suggest that participation calls were not predominant. Only a minority of the tweets related to protest organization and coordination issues. Furthermore, comparing the actual contents of the Twitter information exchange reveals similarities and differences between the three movements, which the different national contexts can explain. The considerable limitation of their methodology is that the authors manually scanned the content of 60 tweets from the Greek and Spanish data sets (30 for each case) to list the specific categories that could demonstrate how Twitter was used for political action.

3.13 Anti-austerity protest in Italy (2011)

On 15 October 2011, around 200,000 people gathered in Rome to protest economic inequality through the political influence of the European Commission, the European Central Bank, and the International Monetary Fund. There were protests across Italy on the same day. Several political parties, trade unions, and civil movements supported the demonstrations, including Cobas, Federazione Anarchica Italiana, Young Communists, Purple People, Workers' Communist Party, Party of Italian Communists, Communist Refoundation Party, Left Ecology Freedom, Critical Left, and many others. Similar to the previous ones, the protests started in solidarity with Spain (Corriere Della Sera, 2011, Milan, 2013).

Vicari (2013) examined the use of social media for public reasoning on social issues. Her research focuses on the 15 October 2011 Twitter stream of the polycentric protest for global change in Italy, the public's use of this medium after large-scale protest activities. Quantitative analysis of over 8,000 tweets shows that Twitter supports public arguments about social conflict. According to Vicari, Twitter is a news platform rather than a conversational platform; its primary purpose is to communicate information available in mainstream media coverage.

3.14 Wisconsin labor protest (2011)

In February 2011, the so-called Wisconsin protests were a series against the Wisconsin Act 10, also known as the „Wisconsin Budget Repair Bill," in February. Protests centered on the Wisconsin State Capitol in Madison, with satellite protests in other Wisconsin cities. In addition, protests were held at the University of Wisconsin–Madison and the University of Wisconsin–Milwaukee. After the Wisconsin Supreme Court upheld the collective bargaining bill on 14 June, protesters dropped to around 1,000. Following the failed recall of Governor Scott Walker in 2012, protests drove state senator recall elections in 2011 and a contentious Wisconsin Supreme Court election in 2011 (Vetterkind, 2021).

In a case study of the 2011 Wisconsin labor protests, Veenstra, Iyer, Hossain, and Park (2014) studied Twitter as a news reporting medium evolving in such crisis circumstances generally. During the first three weeks of the protests, over 775,000 tweets with the #wiunion hashtag were analyzed. Researchers found a substantial variation in tweeting habits between people who use their mobile devices to post to Twitter and those who use their laptops. Fewer links to conventional media and more context are common in the URLs posted by mobile users than on desktops and laptops. As individuals get more adept at providing first-hand accounts of their experiences, the number of links they share drops. The findings of analyses alter if they are restricted to solely original tweets rather than retweeted ones.

Macafee and De Simone's (2012) study investigates the relationship between young people's informational and expressive uses of four social media platforms—Twitter, Facebook, YouTube, and blogs—and offline activism. According to their survey, while college students used these social media platforms to gather information about the Budget Repair Bill protests, only expressive uses were associated with offline protest participation.

3.15 Israel Hamas demonstration, 2011–2012

Palestinians protested in 2011–2012 in the Palestinian National Authority and the Hamas-ruled Gaza Strip as part of the Arab Spring. The protests were against the Palestinian government and in support of the Tunisian, Egyptian, and Syrian uprisings. The primary indicators of the protest were the rising cost of living, the increased fuel prices, and the VAT rate. Massive protests have occurred in Ramallah, Nablus, Balata Camp, Bir Zeit, Jalazun Camp, Hebron, Bethlehem, Beit Jala, Dheisheh Camp, Jenin, Jericho, Tulkarm, and Dura. In addition, road closures, tire burning, self-immolations, peaceful protests, stone-throwing clashes, and worker strikes have characterized 2012 (Bhavnani & Donnay, 2012, Reuters, 2012).

Heemsbergen and Lindgren (2014) investigated the evolving use of social media during combat by examining the Facebook and Twitter profiles of the spokesman for the IDF (Israel Defense Forces). Israel's confrontation with Hamas-affiliated forces in November 2012 provided intriguing data on a sovereign power's wartime use of social media and the resulting networked discourse. They examined Facebook data for effective distribution patterns through content and discourse analysis. In addition, the IDF's social media texts were analyzed using Connected Concept Analysis to identify the construction of meaning. According to the authors, researching this social media data allows the authors to comment on the evolving modes, strategies, and expectations of official public diplomacy, propaganda, and transparency in wartime.

3.16 Brazil Vinegar (2013)

In 2013, demonstrations against the Confederations Cup broke out in several Brazilian cities, sparked by the Movimento Passe Livre (Free Fare Movement), a local advocacy group for free public transportation. The protests began to address government corruption and police brutality against some demonstrators. By mid-June, the movement had grown to its largest since 1992’s anti-Collor de Mello protests. In the Gezi Park protests (Turkey) in 2013, social media has played a crucial role in organizing public outcries and keeping protesters connected.

Bastos, Recuero, and Zago (2014) examined the association between the geographic location of protestors who attended rallies during the protests and those who tweeted about it. They examined the overlap between various sources of geographic information from Twitter, such as geocoding, hashtags, or user profiles, as provided by multiple samples drawn from a population of three million tweets about the events, and compared the data to the location of protestors participating in street demonstrations. The authors corrected for the population’s unequal distribution and conducted geospatial and spatial clustering analyses on sets of spatial locations. Bastos, Recuero, and Zago discovered evidence that people tweeting protests are geographically removed from the mass protests and that users from geographically isolated places depend on Twitter hashtags to participate in the demonstrations remotely.

3.17 OccupyGezi, Turkey (2013)

On 28 May 2013, protests erupted in Turkey, initially against an urban development plan for Istanbul’s Taksim Gezi Park. Outrage over the violent eviction of a sit-in protesting the plan sparked the protests. Following this, protests and strikes erupted across Turkey, focusing on press, expression, and assembly freedom issues and the alleged political Islamist government’s erosion of Turkey’s secularism. Around 3.5 million people (out of a total population of 80 million) have participated in nearly 5,000 protests across Turkey. More than 8,000 people were injured, many critically. Besides the 11 deaths and over 3,000 people were arrested. After the police left Taksim Square on 1 June, the sit-in at Taksim Gezi Park became a protest camp, with thousands of protesters in tents, organizing a library, medical center, food distribution, and media. Some foreign governments and international organizations criticized police brutality and the lack of government dialogue with protesters. The protester’s grievances ranged from local environmental concerns to Recep Tayyip Erdogan’s authoritarianism. Analysts said the protests were the most challenging events of Erdogan’s ten-year term and the most widespread nationwide unrest in decades. The protests have been compared to the Occupy movement and the May 1968 events because there was no centralized leadership (Bee & Chrona, 2017).

Varnali and Gorgulu (2015) contributed to the increasing body of knowledge about online political involvement by examining the social determinants of action that motivate expressive political activity on Twitter. Their findings indicated that social impact variables independently explain a considerable fraction of the diversity in online political participation. Additionally, whereas social identification and group norms were significant predictors of expressive political activity on Twitter, subjective norms had no effect.

3.18 Euromaidan (2013-2014)

The public protests began on 21 November 2013 in Kyiv's Maidan Nezalezhnosti (Independence Square) to lessen ties with Russia and the Eurasian Economic Union. The primary escalator of the protest was that the Ukrainian government chose to suspend the signing of the EU–Ukraine Association Agreement. The protests grew in scope, calling for the resignation of President Viktor Yanukovych and the Second Azarov Government. Intense government corruption, abuse of power, and human rights violations in Ukraine fueled the protests. After the violent dispersal on 30 November, more protesters joined and sparked the Ukrainian Revolution of Dignity. During the Euromaidan, protesters occupied and barricaded the Maidan (central square) in Kyiv and some administrative buildings, including the Kyiv City State Administration. Unrest erupted in January after the Ukrainian parliament passed anti-protest legislation. Across Ukraine, protesters occupied government buildings. Concluding in mid-February, Riot police advanced on Maidan but did not fully occupy it. As a result of these events, Yanukovych and the parliamentary opposition leaders (Vitaly Klitschko, Arseny Yatsenyuk, Oleh Tyahnybok) signed the Agreement on Settlement of Political Crisis in Ukraine on 21 February 2014. Yanukovych and other high-ranking officials fled, then the parliament ousted Yanukovych, installed Oleksandr Turchynov as prime minister, and freed former PM Yulia Tymoshenko (Zelinska, 2017, Oliinyk & Kuzio, 2021).

Lyebyedyev and Makhortykh (2018) examined how social media was used to shape the Euromaidan protests in Ukraine. They examined how the online depiction of protests changed throughout the protest campaign and how the framing of the Euromaidan changed across different language streams by automatically classifying a large amount of Twitter data. Their findings suggest that the social media portrayal of Euromaidan has shifted from a peaceful movement to a revolutionary force and then an existential threat to the Russian-speaking population, fueling the continuation of the political crisis in Ukraine and the annexation of the Crimea accelerated.

3.19 Black Lives Matter (2013)

The Black Lives Matter movement (BLM) is a grassroots political and social movement of a decentralized network of people and NGOs with no formal hierarchy that seeks to raise awareness of racism, discrimination, inequality, and police brutality. The movement and its affiliates usually advocate for necessary policy changes for black liberation. The movement was formed in July 2013 with the hashtag #BlackLivesMatter on social media following

George Zimmerman's acquittal in the February 2012 shooting death of African-American teen Trayvon Martin. The movement gained national recognition following the 2014 deaths of two African Americans, Michael Brown in Ferguson, Missouri, near St. Louis, and Eric Garner in New York City. After 2014, the movement has protested against the deaths of many other African Americans by police or while in police custody. Black Lives Matter activists were involved in the 2016 US presidential election 2015. They grew their project into a national network of over 30 local chapters between 2014 and 2016. After the murder of George Floyd by a Minneapolis cop in 2020, the movement regained national and international attention. Approximately 15 and 26 million people participated in the 2020 Black Lives Matter protests, making it one of the country's most significant movements (Buchanan et al., 2020, Campbell, 2021).

Tillery (2019) used content analysis studies to evaluate how Black Lives Matter social movement organizations use Twitter. The first significant finding was that between 1 December 2015 and 31 October 2016, the modal tweet expressed sadness, fury, or despair over police violence and black deaths. According to his second major conclusion, the Black Lives Matter organizations created more tweets about individual rights than gender, ethnic, and LGBTQ identities. Finally, the report revealed that the activists encouraged their supporters to engage in disruptive political activities less frequently than other activists. Tillery's findings demonstrate that the movement can be understood through resource mobilization and new social movement paradigms.

Ince, Rojas, and Davis (2017) examined how Twitter users interact with BLM through hashtags, altering the movement's framing, which they call "distributed framing." They analyzed 66,159 tweets containing the hashtag #BlackLivesMatter in 2014 when #BlackLivesMatter gained traction on social media. Additionally, they track the other hashtags alongside #BlackLivesMatter to determine how online communities impact the movement's framing. They discovered that #BlackLivesMatter is linked to five distinct types of hashtags. These hashtags indicate solidarity with or endorsement of the movement, reference police violence, discuss movement tactics, reference Ferguson, or express anti-movement sentiments.

Edrington and Lee (2018) emphasized that while public relations research has gradually incorporated the study of advocacy organizations, little research has focused on social movements in particular. Through a content analysis of all public tweets sent by

Black Lives Matter over four years, their study examined the message strategies used on Twitter by the social movement to share information, build community, and promote action. Consistent with research on other types of organizations, informational messages proved to be the most common. However, the study also analyzed these strategies' influence on audience engagement in terms of replies and retweets. The findings of Edrington and Lee suggest that community-building messages garner the most retweets but found no significant differences in terms of reactions.

Wilkins, Livingstone, and Levine's (2019) research investigated the rhetorical roles of Twitter use during a formative period of the Black Lives Matter social movement. They analyzed how activists utilized Twitter to balance competing social change goals, such as expanding the movement beyond members of disadvantaged groups while avoiding appropriation of their message by „allies" from advantaged groups. They discovered that while Twitter users support many and frequently conflicting conceptions of the issues represented by the movement, rhetorical methods are utilized to promote inclusive meanings that emphasize racism. Additionally, when activists address alternate definitions of movement players and concerns, they employ representations of otherness to represent supporters of these definitions as anti-movement. Finally, the researchers discovered that one strategy to resolve the tension between movement growth and fostering disadvantaged-group control is to use identity and technology resources to identify how diverse groups might be movement advocates and social change action tactics.

3.20 Romanian protests (2017–2019)

In January 2017, various anti-government and corruption protests started in Romania against the planned ordinance legislation proposing the pardoning of certain offenses and amending the Romanian Penal Code. Despite public and judicial opposition, the newly sworn-in administration quietly approved an ordinance amending the Penal Code and Penal Procedure Code on 31 January. Decriminalization of government corruption, opponents claimed, will help hundreds of present and past politicians avoid ongoing criminal investigations or prison sentences. Under the organization of the Corruption Kills (CU) community, more than 37,000 people opposed that night after the ordinance was passed. Less than a week later, on 5 February, over 500,000 Romanians demonstrated across the country, making it the largest protest since the fall of Communism. Since the major grievance of the demonstrators was not addressed, the rallies continued practically daily throughout the country, with more and more protesters demanding early elections and the administration's resignation. After the winter of 2017, 50,000–100,000 Romanians took to the streets on 20 January 2018 to protest proposed changes to the penal code and the justice system regulations. While smaller-scale protests continued daily, large-scale protests erupted in Bucharest on 10 August 2018, with the slogan "Diaspora at Home." Protesters pushed the government to remove the decree in 2017, and Florin Iordache, the justice minister who put it forward, resigned shortly after due to the scandal (Ilie, 2017, Roberts & Stan, 2017, Alexe, 2018).

Mercea (2020) examined transnational expatriate action through the lens of open-accessible social media pages. His study the links between users of Facebook event pages affiliated with 122 places globally where demonstrations in favor of Romania's anti-corruption #rezist protests. Examining the relationships between sociodemographics, spatial, and network characteristics revealed a connectedness disparity between geographic location and gender of page users in comment and share networks. When individuals were engaged on the same pages, connections rose, although users' shared location was associated with substantially lower posting degrees but not sharing activity. Simultaneously, although more of them were male, users exhibited a systematic inclination to interact with the opposite gender. According to Mercea, thus, renewed emphasis should be devoted to socio-spatial disparities in using social media platforms that localize and connect transnational activism.

Mercea, Burean, and Proteasa (2020) investigated the use of public Facebook event pages and the effect of political information on student involvement. Additionally, they investigated whether political material was available on protest-related pages. Additionally, the researchers examined the social network structure formed by those pages to comprehend its spread within that public domain. They discovered that political information constituted a significant component of open, although localized, activist engagement on Facebook, with students who followed a page being more likely to participate in demonstrations. According to the researchers, these findings allow for evidence-based discussion of the relationship between individual demand and the supply of political information on social media and protest involvement.

3.21 Serbian protests (2018–2020)

In late 2018, Serbia and President Aleksandar Vučić (Serbian Progressive Party, SNS) faced emerging but mostly peaceful anti-government protesters who were stimulated by an attack of the leaders of the opposition coalition Alliance for Serbia. On 23 November 2018, attackers injured Borko Stefanović (Serbian Left Party, LS) with steel rods. The protests, which were further heated by many scandals involving ruling party members, have lasted more than a year, making them the most extended anti-government demonstrations in Serbia. Parallel to the protests, Vučić and SNS organized rallies in all Serbian districts, while pro-government media attacked demonstrators and opposition leaders, tying them to Nazism and distributing misinformation to their readers.

Although there is no specific protest literature about this event, Twitter announced a takedown of approximately 8,500 accounts and more than 43 million tweets targeting Serbian Twitter users on 26 March 2020. According to the analytics of Stanford Internet Observatory, these accounts and tweets acted in concert to promote President Vučić and the SNS and attack his opponents. According to the analytics, the eliminated accounts worked for several years—increasing their action in mid-2018 and 2019.—to promote the SNS and Aleksandar Vučić. They accomplished this by retweeting and responding to Vučić-aligned tweets (more than 12.5 million retweets from pro-Vučić accounts); sharing links to content on Vui-aligned websites (more than 14.8 million links); and attacking Vučić's opponents, particularly the Alliance for Serbia. Although, according to Bush, a direct link between this network and SNS has not been formed, there can be no doubt that, based on the content shared by these accounts and the time period during which they were active, this network was created to boost Vučić's election chances in early 2017 (Bush, 2020). Overall this report shows the role and possibility of fake accounts in the distortion of information.

3.22 Hong Kong Anti-Extradition Law Amendment Bill Movement protests (2019–20)

The movement began on 15 March 2019 in response to the Hong Kong government's introduction of the Fugitive Offenders Amendment Bill regarding extradition. The demonstrations began with a sit-in protest at the government headquarters, followed by a massive rally in June outside the Legislative Council Complex to stop the bill's second reading. On 16 June, a day after the Hong Kong government postponed the law, a more significant rally demanded its abolition. Activists made five demands throughout the protests. On 4 September, the bill was withdrawn, but demands were refused, which further escalated the protest. The activists conducted sieges of two universities and the storming of the Legislative Council. During the protests, hundred people were arrested, including notable campaigners (Barron et al., 2019, Tang & Cheng, 2022).

Qi, Jiang, Bu, and colleagues (2019) investigated English-language social interactions regarding the Hong Kong Protests, a sequence of events that dominated social media in 2019. Their system scraped Twitter and Reddit for data via their APIs. They conducted sentiment analysis, and the results indicated changes in public sentiment after significant events. Influencers - users at the core of social dialogues - were identified through social network analysis. Their interactive Tableau dashboard enables users to effortlessly monitor live English-language social media conversations concerning Hong Kong protests.

Table 2: Summary of the protests (2009-2019)

Name	Date	Location	Goals	Methods	Resulted in
Moldovan parliamentary election protests and civil unrest (2009)	6 – 12 April 2009	Chișinău, Cahul, Orhei, Bălți 13 cities in Romania	Election transparency, recount of the votes, a new election, or resignation of the government	Demonstrations, riots	Partly successful, the parliament failed to elect a new president
Student protests in Austria in 2009	20 October 2009 – 14 May 2010	Vienna, Graz, Linz, Salzburg, Innsbruck (Austria)	Democratization of the universities, abolition of tuition fees	Occupation	Resulted no crucial changes in the Education policy
Iranian presidential election protests 2009-2010	13 June 2009 – 7 December 2010	Iran (multiple cities)	Election transparency, recall of Mahmoud Ahmadinejad from office	Demonstrations, riots, civil disobedience, strike actions	None, protests suppressed by force
Tunisian revolution (2010-2011)	17 December 2010 – 14 January 2011	Tunisia	Election transparency, democratisation	Civil resistance Demonstrations General strikes Self-immolations Spontaneous uprisings	Overthrow of the government, resignation of Prime Minister, dissolution of the political police, release of political prisoners, new elections
Venezuela (2000s)	Late 2000s and early 2010s - ongoing	Venezuela	Solve socioeconomic and political crisis	Demonstrations, riots, civil disobedience, strike actions	Ongoing
Protest against Stuttgart 21 (2010)	1 October 2010	Stuttgart, Augsburg (Germany)	Stop the development project	Demonstrations, petitioning, riots	Partly successful, referendum was held, made Green politics stonger
G20 Toronto summit protests (2009)	18–28 June, 2010	Toronto, Ontario, Canada	Fight for social equality	rally, demonstration, rioting, vandalism	Ongoing
Egyptian revolution (2011)	25 January – 11 Feb, 2011	Egypt, Cairo	Democratisation	Civil disobedience; Civil resistance;	Overthrow of Hosni Mubarak after 18 days of demonstrations.

				Demonstrations; Online activism; Riots; Self-immolation; Strike actions	
England riots (2011)	6–11 August 2011	boroughs of London		Rioting, looting, murder, protest march	None
Occupy Wall Street (2011)	17 September 2011	Zuccotti Park, New York City	Reducing the influence of corporations on politics, forgiveness of student loan debt, bank reform, balanced distribution of income	Occupation, demonstration, civil disobedience	No significant result
15-M , Indignados (2011)	15 May 2011 – 2015	Spain	Direct democracy, reduce influence of economic powers in politics	Demonstrations, civil disobedience, civil resistance, rioting, sit-ins, online activism, occupations	No significant result
Aganaktismenoi movement, Greece (2011)	5 May 2010 – 18 October 2012	Greece		Demonstrations, strike action, sit-ins, occupations, civil disobedience	No significant result
Italy (2011)	15 October 2011	Rome, Italy	Solve socioeconomic problems	Demonstations, riots	
Wisconsin labor protest (2011)	14 February 2011 – 16 June 2011	Madison, Wisconsin, USA	Revoke Budget Repair Bill	Protests, sit-ins, demonstrations, recall elections, quorum-busting	None, Budget Repair Bill passed
Israel Hamas demonstration 2011–2012 Palestinian protests (2012)	3 February 2011 - 2 October 2012	Palestine, Gaza	Resignation of President Mahmoud Abbas	Demonstations, riots, self-immolations	None, protests suppressed by force

Brazil Vinegar (2013)	April – July 2013	Brazilia	Reduce public transportatio fee	Occupations, demonstrations, riots	Improved public transportation
OccupyGezi, Turkey (2013)	28 May 2013 – 20 August 2013	90 locations in Turkey	Resignation of government, Protecting Gezi Park	Occupations, riots, protest marches, civil disobedience, petitioning	Successful, plans were cancelled
Euromaidan (2013-2014)	21 November 2013 – 22 February 2014	Kiev, Ukraine	Signing of the EU Association Agreement and Free Trade Agreement, Impeachment of President Viktor Yanukovych	Demonstations, civil disobedience, occupation	Successful, Removal of Viktor Yanukovych from office, Former prime minister Yulia Tymoshenko freed from jail, The new Government of Ukraine resumed preparations in signing of the EU Association treaty.
Black Lives Matter (2013)	2013 – ongoing	International	Dismantling white supremacy	Demonstrations, riots	
Romanian protests (2017–2019)	18 January 2017 – 10 August 2019	Romania	Modify criminal code provisions, resignation of the government	Demonstrations, protest marches, sit-ins, occupations	Withdrawal of the decrees, Resignation of the Minister of Justice
Serbian protests (2018–2020)	7 July 2020 – ongoing	Serbia	New parliamentary elections, Release of arrested protesters, Release of arrested protesters	Demonstrations, civil disobedience	Partly successful, 10 protesters released, reintroduction of curfew cancelled
Venezuelan protests (2019)	10 January, 2019 – 16 November 2019	Venezuela	Resignation of Nicolas Maduro, Free elections	Demonstrations, riots	None
Hong Kong Anti-Extradition Law Amendment Bill Movement protests (2019–20)	15 March, 2019	Hong Kong	Full withdrawal of the extradition bill	Demonstrations, riots	Government crackdown

The previous pages provided a detailed overview of various social media-mediated protests and the related case studies from the last decade. The first chapter of this work helped us place these demonstrations in the changed circumstances of the internet age. In addition, the effectiveness of the previous protests is understandable if we consider the last three-part classification. Crowd-enabled collective networks alone, like the riots in England that followed police brutality against Mark Duggan, have not been able to make significant advances or changes in the system. However, organizationally brokered class-action lawsuits already have the infrastructure to promote and support a case, most often meeting the needs of a significant portion of the population of a given area. Thinking of the G20 protests or the Occupy Wall Street movement is enough. However, it was also apparent from the previous examples that the most effective organizational type was hybrid connective action simply because it combined the growing excitement and need for crowd-enabled action with resource-rich organizations. For instance, in Moldova, a small part of the population initially expressed the wish for new elections to be broadcast to a broader population of Moldova. During the Unibrennt protest in Austria, student organizations were able to build a nationwide protest and occupation movement centered on the intense needs of university students. The previous pages also revealed that the ongoing protests (e.g., Black Lives Matter) showed an evolution from a crowd-enabled demonstration to Hybrid Connective Action, increasing its strength and impact over the last few years became apparent; the protests developed significantly after the death of George Floyd.

In the last decade, the use of social media data as a source of information has become increasingly widespread in various protest case studies. Although the social media-based analytical methods of collective actions have accelerated the slow and expensive conventional methods, such as surveys, they are still characterized by various deficiencies, such as lack of representativeness or transferability of the results. Moreover, the widespread and well-established advantages of the network modeling approach on hashtags (e.g., Venezuela, Stuttgart 21, G20 above) or users (e.g., Egypt, Occupy Wall St.) in recent literature analyzing protests based on social media data are often not capable of handling the complexity of the sentiment, temporal, and spatial patterns of these actions in one approach. One limitation is that the analysis relies on Tweets having coordinates as an inherent part of the dataset (Felmlee et al., 2020, Howard et al., 2011), which—according to earlier studies—represents only a tiny subset of all tweets (approximately 1–10%)

posted within a specified period (Morstatter et al., 2013). Another limitation is that they focus only on a single language (usually English) which may also limit the spatial interpretability of the results (Drüeke & Zobl, 2016), especially in the case of movements that span multiple countries. This is where the benefits of a comprehensive data pre-processing method, including translations and user location extraction, become relevant, which supports the more precise assessment of a post-event situation by extracting an additional information layer from users' digital footprints by revealing contextual information insights. This approach already exists in other fields, for example, to uncover disaster footprints, but not in the case of social movements or unrest (e.g., Resch et al., 2018, Wang et al., 2016, Ye & Wei, 2019).

RESEARCH OBJECTIVES AND RESEARCH QUESTIONS

This dissertation suggests a comprehensive methodology to overcome the limitations mentioned above in the existing methods and handle the complexity of protest analyses. The proposed approach includes multi-lingual corpus translation and location and sentiment extraction, using machine-learning topic modeling methods to reveal collective action's hidden interests and motivators. Through this, our approach has a distinct advantage over the prior investigations that primarily focused either on hashtag activism (LeFebvre & Armstrong, 2018, Sinpeng, 2021) (ignoring the spatial dimensions) or, on the contrary, using only location-specific hashtags (Conover et al., 2013, Croeser & Highfield, 2014). In contrast, by applying machine learning algorithms and techniques that are almost entirely automatable we can have a much more comprehensive range of input data than in existing studies, where the researchers solely evaluate posts manually (Panagiotopoulos et al., 2014, Maireder & Schwarzenegger, 2012).

This work examines the similarities and correlations of Twitter data's spatial, temporal, and sentimental markers by developing a new data-driven combined approach to investigating two distinct East European protest movements. First is the European influence of Slovakian journalist Ján Kuciak and his fiancée Martina Kusnirova's assassination in 2018. The murder of Kuciak and the people's reaction were recently analyzed in a broader timeframe (28 February–28 July 2018) by Kapanova and Stoykova (Dimov & Fidanova, 2021). The authors applied a network-based analysis to the #AllforJan hashtags (Dimov & Fidanova, 2021). Second is the influence of the Belarusian presidential election in 2020 (9 August–23 September 2020), which was not analyzed by contemporary research.

The multi-spectral interpretation of the dynamic and challenging nature of the events requires an efficient analytical method to assist in a more comprehensive understanding of the protest dynamics. Thus, our work goes beyond state of the art in two distinct ways:

1. We demonstrate how georeferenced social media data can be used for analyzing political events, even at a smaller spatial and societal scale and in unique non-English languages.

2. From a methodological viewpoint, the present work proposes a new algorithmic workflow that combines time-series clustering with semantic topic modeling and sentiment analyses on georeferenced social media data.

This research will put effort into profoundly understanding the relevant research of the past decade. By presenting an original perspective and considering the presented limitations of existing analyses, in this dissertation we intend to answer the following research questions:

1. Temporal and spatial aspects:
    1. How tweeting activity related to the murder of Kuciak varied over time throughout Europe? (RQ1a)
    2. How tweeting activity related to the presidential election of Belarus varied over time throughout Europe? (RQ1b)
    3. Can we identify the influence of specific events and incidents, such as media reports or findings of the investigation based on this tweeting activity? (RQ1c)

2. Content aspects:
    1. How does the sentiment of the tweets vary over time, and how does it relate to specific events and news? (RQ2a)
    2. Around what topics do the tweets revolve, beyond the murder itself? (RQ2b)

3. Profiles:
    1. How can we characterize the countries based on the temporal aspects of the tweeting activity and the sentiment of the tweets? (RQ3a)
    2. Does the categorization of the countries also reflect differences in the identified topics or the changes of the sentiment values over time? (RQ3b)

Not alone, but in stack, these questions help to identifies the most similar countries in a space-time cube and separating them into distinct clusters whose members share similar time series characteristics. Then we are able to analyze the existence and correlations of different constructing logics: collective or connective or perhaps both, among the topics.

METHODOLOGY

4 Twitter

The social media data analyzed in this dissertation were obtained using the Twitter Streaming Application Programming Interface (API), the US-based social networking and microblogging service. The platform was founded in March 2006 and launched five months later by Noah Glass, Jack Dorsey, Biz Stone, and Evan Williams (Twitter Blog, 2012). By 2012, the service had grown to over 100 million users and handled an average of 1.6 billion search queries per day. It was one of the top ten most visited websites in 2013 (D'Monte, 2013). Twitter had a monthly active user base of over 330 million in Q1 2019 (Molina, 2017). It is commonly used by politicians, journalists, and other public figures, making Twitter the perfect platform for spotting trends, news, and impact.

4.1 The Twitter data

The Twitter API provides an advanced programmatic interface for the site's database to retrieve and study core elements like posted messages called "tweets," direct messages, and users as well as their attributes. The Twitter API has evolved, adding additional layers of access, among others, for academic researchers to enhance public discourse and research. Although Twitter recently released the 2nd version of the Twitter API, enhancing features and a simplified onboarding process, this dissertation's data relies on the first version of Twitter API. The subsequent paragraphs will describe this version's access level and services of the Twitter API, focusing on the two essential parts of our analysis: the tweets and the user and well as their attributes (Twitter Inc., 2018).

Accoring to the most recent data, Twitter is the second most popular social networking site worldwide ranked by a number of monthly active users. In context, nodays Twitter has 7.91 percent of the global social media user population. According to the data of Statcounter, after the Facebook, the Twitter was the second most popular social media site in the countries of Europe, including the analyzed coutries and timeframe of this thesis (Statcounter 2018-2020).

Due to the fact that Twitter was first developed in the United States, the majority of the site's users are also American. Despite this, Twitter is quickly penetrating markets in other countries throughout the world. In point of fact, the most of Twitter's recent expansion has taken place in countries other than the United States. According to the statistics, the majority of Twitter users are male: 68.1 percent men, 31.9 percent are female. While other social networking sites, such as Instagram and LinkedIn, have a more even gender ratio, with 51% and 43% of members being female, respectively. Twitter's readership consists mostly of younger and middle-aged individuals. In actuality, 59.2% of users are between the ages of 25 and 49. Consequently, the age group of the main user base of Twitter correlates with the age of protesters that we have seen in the case-studies of the literature review.

Table 3: Age distribution of Twitter users

Age group	Share (%)
13-17	6,8%
18-24	17,1%
25-34	38,5%
35-49	20,7%
50+	17,1%

4.2 The Tweet

The ground subset of our analytics is the tweet, a message posted on the sender's profile page that may contain text, photos, a GIF, and video. Other users can forward individual tweets to their feed, known as a "retweeting" process. In 2015, Twitter launched "quote tweet," allowing users to add a comment to their retweet. Users can also "like" individual tweets. The counters for "likes," "retweets," and replies appear next to the respective buttons in timelines, such as on profile pages and search results on the web and mobile interface. Initially, Twitter began as an SMS-based service that limited the length of a tweet to 140 characters. Since 7th November 2017, the current character limit for a tweet is 280 characters. However, some Unicode glyphs cont as more than one character; among these are the emojis (pictograms) embedded in text and count as two characters. Chinese, Japanese and Korean languages also count as two characters, so tweets composed in these languages can only have a maximum of 140 character-length. In addition, tweets may have entity objects that impact the message's length (Twitter Inc., 2018).

4.3 Tweet Data Dictionary

The retrieved data belong to two main groups: Tweet-specific and User-specific data. The following tweet data dictionary contains the main attributes of a tweet. The following tables are derived from the Twitter Developer Platform; however, they are reduced according to the data types used in the present research.

Table 4: Description of Tweet-specific data (Twitter Inc., 2018)

Attribute	**Type**	**Description**
created_at	str	UTC time when this Tweet was created.
id	int; str	The integer representation of the unique identifier for this Tweet. This number is greater than 53 bits and some programming languages may have difficulty/silent defects in interpreting it. The string representation of the unique identifier for this Tweet. Implementations should use this rather than the large integer in id.
text	str	The actual UTF-8 text of the status update (tweet). Tweets are publicly visible by default, but senders can restrict message delivery to only their followers.
source	str	Utility used to post the Tweet, as an HTML-formatted string. Tweets from the Twitter website have a source value of web. (E.g., "Twitter Web Client")
coordinates	coordinates	Nullable. Represents the geographic location of this Tweet as reported by the user or client application. The inner coordinates array is formatted as geoJSON (longitude first,

		then latitude)
place	places	*Nullable* When present, indicates that the tweet is associated (but not necessarily originating from) a Place
retweet_count	int	Number of times this Tweet has been retweeted. A particular type of Tweet is a re-posted Tweet, the Retweet. This feature helps users and organizations to share a specific Tweet with all of their followers quickly. Users can Retweet their Tweets or Tweets from someone else. Sometimes users type "RT" at the beginning of a Tweet to indicate that they are re-posting someone else's content. Research tends to ignore Retweets from analytics because they are not unique products of a user.
entities	entities	Entities which have been parsed out of the text of the Tweet. This research uses three of them: url, hashtag and user_mentions. URL: Uniform Resource Locators (URLs) in a Tweet is shortened by Twitter that uses t.co abbreviation, which hides the original URL. The current length in a Tweet is 23 characters, even if the length of the URL would be shorter. Hashtag: A hashtag is a metadata tag that is prefaced by the pound sign or hash symbol, #. Hashtags are used as a form of user-generated tagging that enables cross-referencing of content; that is, sharing a topic or theme. For example, a search within Instagram for the hashtag #review returns all posts that have been tagged with that hashtag. user_mentions: A user_mention is a metadata tag that is prefaced by the at sign, @. The entities section will contain a user_mentions array containing an object for every user mention included in the Tweet body, and include an empty array if no user mention is present. It is also a reply is a response to another Tweet, and is is enable users to join in a conversation as it is happening on the site. In the case of a reply Tweet, @names appears at the

		beginning of a reply Tweet and it will not count towards the character limit. New non-reply Tweets that begin with a @mention, as well as @mentions explicitly added by the user in the text of the Tweet, will count.
favorited	bool	*Nullable.* Indicates whether this Tweet has been liked by the authenticating user.
retweeted	bool	Indicates whether this Tweet has been Retweeted by the authenticating user.
lang	str	Nullable. When present, indicates a BCP 47 language identifier corresponding to the machine-detected language of the Tweet text, or und if no language could be detected.

4.4 User data dictionary

The following user data dictionary contains the main attributes of a user. Registered Twitter users can post, like, and retweet tweets. They interact with Twitter through browser or mobile frontend software or programmatically via its APIs. The user data dictionary contains the essential attributes of a user's activity.

Table 5: Description of user-specific data (Twitter Inc., 2018)

Attribute	Type	Description
id; id_str	int; str	The integer representation of the unique identifier for this User. This number is greater than 53 bits and some programming languages may have difficulty/silent defects in interpreting it. Using a signed 64 bit integer for storing this identifier is safe. The string representation of the unique identifier for this User. Implementations should use this rather than the large, possibly un-consumable integer in id.
name	str	The string representation of the name of the user, as they’ve defined it. Not necessarily a person’s name.

		Consequently it is not equal the screen name, handle, or alias that this user identifies themselves with. screen_names are unique but subject to change. Use id_str as a user identifier whenever possible. Typically a maximum of 15 characters long, but some historical accounts may exist with longer names.
location	str	The string representation of the user-defined location for this account's profile. Not necessarily a location, nor machine-parseable. This field will occasionally be fuzzily interpreted by the Search service.
description	str	Nullable . The user-defined UTF-8 string describing their account. It is a short public resume about users or organizations shown beneath the Twitter profile picture. The bio field can include a maximum of 160 characters of text, hashtags, emojis, and link of profiles with which users or organizations are affiliated.
friends_count	int	The number of users this account is following (AKA their "followings"). Under certain conditions of duress, this field will temporarily indicate "0".
statuses_count	int	The number of Tweets (including retweets) issued by the user.
created_at	str	The UTC datetime that the user account was created on Twitter.

4.4.1 Location fields

Although the tables above offer a specific overview of different data attributes, various location-specific data types appear multiple times that require additional explanation. Similar to the dictionaries above, Twitter data contains two main classes of geographical metadata, (1) the Tweet location and (2) the Account location. The Tweet location is available when a user shares location at the Tweet message, while the Account location is based on the so-called home location provided by the user in their public profile. The significant difference between these geographical classes is that the account location is a free-form 30 character length field and may or may not hold metadata that can be geo-referenced (Twitter Inc., 2018).

4.4.2 Tweet location

Twitter offers multiple ways to filter messages and posts by specific location data through its various operators. Twitter allows users to add the exact location of where they were tweeting from using coordinates. Tweet-specific location information falls into two general categories: Tweets with a specific latitude/longitude “Point” coordinate Tweets with a Twitter “Place.”

The first category, tweets, contain precise latitude/longitude coordinates representing a point somewhere in the world, and this information comes from the built-in GPS receiver of the device. This location type does not include further information beyond the coordinates, such as from which city or country it was precisely posted. The other location type that the user can assign to a tweet is the Twitter “Place” tag. In contrast with adding only coordinates, this tag has different properties, such as the name of the city or region and the country code showing the country where the given “Place” is located. In addition, it contains a polygon consisting of 4 longitude-latitude coordinates representing the general area where the user is posting the Tweet message.

Additionally, the “Place” will have a display name, type (e.g., city, area, neighborhood), and the country code corresponding to the place’s country, among other fields. The source of this information is still the built-in GPS (or GNSS) receiver of the device, but how this information is visualized and presented is different from the option written above, where exact coordinates were attached to a tweet. Twitter provides geographical coordinates in longitude, latitude, and, consequently, reverse order than the custom. However, the Twitter PowerTrack Geo-Operators expect coordinates in the longitude, latitude order. Although Twitter has multiple location filtering options below we will discuss those that were used in this research: The point_radius operator and the bounding_box.

Twitter's point_radius operator allows defining a circular geographic area and customizing posted Tweets with tweet-specific location data within that zone. The point_radius operator takes a central point with longitude-latitude coordinates and a radius of up to 25 miles (40.2 km). Any tweet with a geopoint that falls within this region will be matched. Additionally, tweets containing Twitter locations will match where the geographic polygon specified for the location falls entirely within the defined point radius range, while Places whose polygons are outside of the specified point radius zone will not match.

The bounding_box operator allows defining a 4-sided geographic area and customizing posted Tweets with tweet-specific location data within that zone. The bounding_box operator takes opposite corners of the box with longitude-latitude coordinates, where each side of the box is up to 25 miles in length. Any tweet that falls within this region will be matched. Additionally, tweets containing Twitter locations will match where the geographic polygon specified for the location falls entirely within the defined point radius range, while Places whose polygons are outside of the specified point radius zone will not match. It shall also be noted that the tweet and user description field may contain further geographical data that requires additional geoparsing methods (Twitter Inc., 2018).

5 Data

5.1 General considerations of data handling

Generally, user privacy is an essential question of investigations that rely on individuals' social media data. In this relation, it shall be noted that we do not use and communicate identifiable user data; in fact, our research focuses on the big scale and similar tendencies of various countries. Thereby individuals' private data has no direct reflection in the outputs, and throughout the whole analysis, it is only used to group users at the country level. Moreover, we were analyzing collective action, and most of the tweets were published in response to the murder, so we interpreted them directly and did not infer any further information that was not the original motivation of someone for posting a tweet.

5.2 #Allforjan (2018) dataset

The first dataset were obtained using the Twitter API (Twitter Inc., 2018) for the period between 26 February and 15 March 2018. The starting date is adjusted to the first official report of the murder of Ján Kuciak, while the final day is adapted to the earliest statement of the resignation of Prime Minister Robert Fico. The dataset consists of the content of the tweets and additional attributes such as user name, user location, and the timestamp when the tweet was posted. The first query used multiple bounding boxes that overlapped the official boundaries of Slovakia. This methodology resulted in approximately 45 000 tweets from the timeframe mentioned above; however, only a few tweets contained topic-specific data. In the second round, we harvested tweets with relevant content (such as names: kuciak, kusnirova, fico) or hashtags (#AllforJan) within this period. This resulted in 13,176 topic-specific tweets from all over the world. However, as our analysis approach requires geospatial analyses as well, we have implemented secondary filtering on these datasets to identify relevant tweets by focusing on place attributes or the coordinates of a tweet, out of which at least one parameter should contain valuable data. Around 3000 tweets where user location was not specific enough (such as “World,” “Internet,” or “online”) have been excluded. Most of the tweets were posted from Europe; thereby, we consider only European locations for the rest of the analysis. By transforming user location into coordinates, we could further increase the number of tweets used for the analysis, as the original dataset contained only 24 tweets with coordinates. Another 1800 tweets were also removed from our dataset as they were too short or no meaningful coordinate could have been attached to them. Overall, this two-step query and the following filtering resulted in 8069 Tweets distributed over an 18-days long timeframe for 39 countries (see Figure 1 for an overview of the timeline).

5.3 #Belarusprotest (2020) dataset

The data for the second dataset was acquired using the Twitter Streaming Application Programming Interface (Twitter Inc. 2018) between 9 August and 23 September 2020. The starting date is adjusted to the day of the Belarus presidential election, while the final day is adapted to the formal inauguration of the president of Belarus, Alexander Lukashenko.

The dataset includes the tweets' content and additional attributes such as the user's name, location, and the time the tweet was posted. The analyzed dataset was built on two queries in the same time frame.

The first dataset contains geolocated tweets from the five most populous administrative centers of Belarus: Minsk (1.8m), Homyel (489,000), Mogilev (361,000), Vitebsk (354,000), Hrodna (339,000). We used the Twitter API's point_radius operator for the query, which allows us to specify a circular geographic area (up to 25 miles) and match messages containing specific location data that fall within that area. Any tweet containing a geographic point within that region will be matched. Additionally, tweets containing Twitter locations match where the geographic polygon defined for the location falls within the defined point radius range. This query resulted in 46,277 geolocated tweets from Belarus.

In the second round, we harvested tweets with relevant content (such as names: belarus, lukashenko) or hashtags within this period. Although collecting and processing topic-specific tweets goes beyond the current state of the art, we should also note that this approach may be limited by different transliteration styles (Lukashenko, Lukashenka, Lukaschenka, Loukachenko). Efforts have been made to overcome this limitation by using name transliteration of the ten most-spoken languages globally that represent more than a 60% of the world population. This resulted in more than 1 million topic-specific tweets from all over the world. However, as our analysis approach focuses on user classification besides geospatial analyses, we have implemented secondary filtering on the datasets to identify and exclude irrelevant tweets by focusing on advertising, spam, and weather accounts. Other tweets were also removed from our dataset as they were too short, or no meaningful coordinate could have been attached to them. This two-step query and the following filtering resulted in 949,321 topic-specific Tweets over a 44-days long timeframe worldwide. Overall the final dataset of Belarus contains 995,598 Tweets.

6 Data Pre-Processing

Our pre-processing approach (Step 1 performed on the raw tweets) consists of a comprehensive text cleaning workflow and transforming available meaningful location information to coordinates for further utilization in the spatial analysis step (Figure 1). In the first part of the pre-processing, we implement primary filtering on our dataset to ignore short tweets that hardly bear any semantic significance. Moreover, stop-words, rare, and too frequent words are removed to normalize the dataset and reduce the redundant noise from tweets (Steiger et al., 2016). Then, we use the available location information of the user to localize the tweets without direct coordinates attached to a tweet (also called geoparsing, see further details in Section V.4) to further increase the size of the dataset for the machine learning-based translation and spatiotemporal analysis.

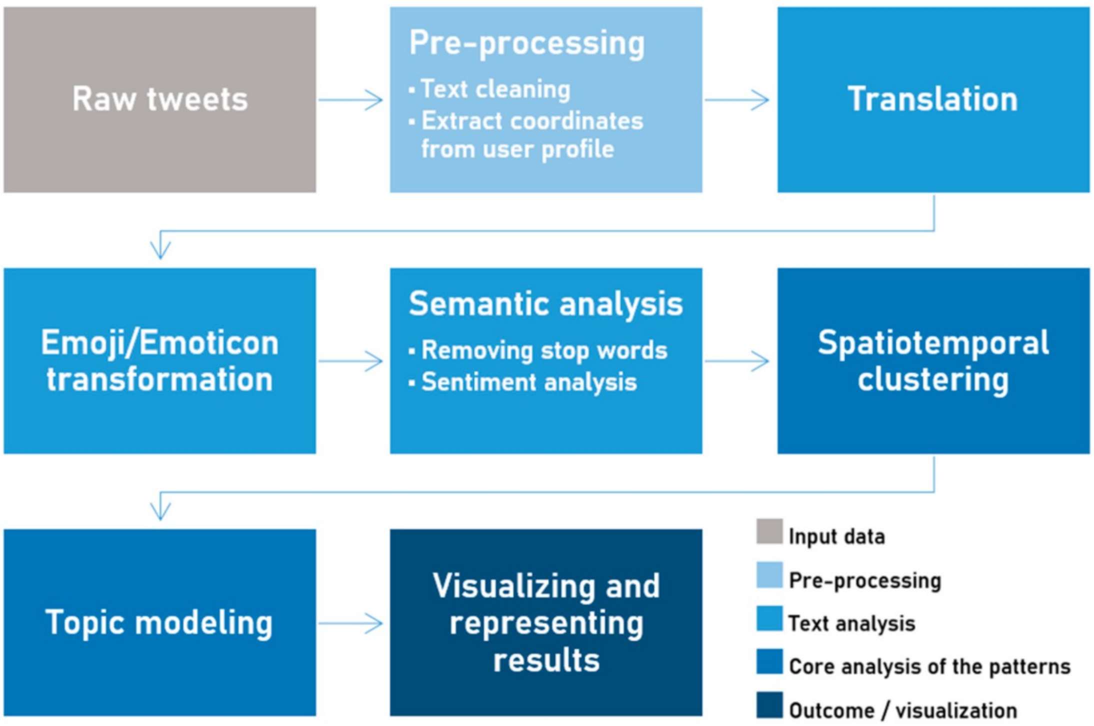


Figure 1: The key steps of our analysis performed on the datasets

6.1 Text Cleaning

This step aims to increase the efficiency of the subsequent translation process. First, we remove short tweets (containing a single word) or those posts that contain only hashtags or URLs because they hold unclear and hardly interpretable semantic value (Pak & Paroubek, 2010). Then, we remove replies (@user_name) from the text as that is considered unnecessary noise in the analysis; in contrast, the hashtags are preserved, but their sign (#)

is removed. The consideration behind this step is that users tend to use hashtags as an integral part of the syntax (Bansal et al., 2015), e.g., “on Friday I am also going to #Bratislava to protest” -> “on Friday I am also going to Bratislava to protest.” As a final cleaning step, we remove newline characters and additional whitespaces from the tweets.

6.2 Locating Tweets Using Coordinates and User Profile Information

As our research questions heavily rely on spatial information, we attempted to process information from all Twitter fields that may contain relevant spatial attributes to increase the number of tweets having an identifiable location, at least at the country level. However, in general, the tweets that inherently include coordinates constitute only a tiny subset of all tweets. Our analytical approach tried to locate those tweets with no coordinates using location information available in the users’ profiles to overcome this limitation.

Twitter allows users to add the exact location of where they were tweeting from using coordinates. Such tweets contain precise latitude/longitude coordinates representing a point somewhere in the world, and this information comes from the built-in GPS receiver of the device. This location type does not include further information beyond the coordinates, such as from which city or country it was precisely posted. To obtain address information for these coordinates, we used the Geopy Python client to access the geocoding web services provided by the OpenStreetMap API (Nominatim). The location.raw (‘address’) function returns a dictionary of address components, such as country code, city, or road, allowing for a targeted query of relevant address information. The other location type, which the user can assign to a tweet (the only way to add location information when this paper was written) is the so-called Twitter “Place” tag. In contrast to adding only coordinates, this tag has different properties such as the name of the city or region and the country code showing the country where the given “Place” is located. The source of this information is still the built-in GPS (or GNSS) receiver of the device, but how this information is visualized and presented is different from the option written above, where exact coordinates were attached to a tweet.

To increase the number of tweets having some spatial reference, we can extract location information from the profile of the Twitter users, which information can serve as a proxy to where this user might be active most of the time. Several data fields fall into this category, but all represent information that the users insert at the account level and not for each tweet separately; thus, the information's credibility mainly relies on the user. Moreover, even if the location information at the account level is valid, it might not be correct for each tweet, for example, if the person is traveling abroad. Generally, these values are not frequently altered and do not necessarily describe the tweet's exact location, but they may represent the user's residence, at least on a city level. However, as our research considers tweets aggregated on a country level and not the exact location within a city or a country, we still find these data valuable as a proxy in instances where no direct location information for a given tweet was provided. To obtain useful information on user_location data, we first ranked individual user_location data by frequency then used the built-in map function of Python and a translator dictionary developed by us to transform all location items and similar group entries. For instance, "BaWü, DE" was transformed to "Baden-Württemberg, Germany" and similarly "B. Württemberg" was also converted to "Baden-Württemberg, Germany." The second step after user_location transformation was to apply the Geolocator function of OpenStreetMap through Geopy, which provided latitude and longitude coordinates that we will use for subsequent mapping applications.

6.2.1 Account Location and User Description

On the other hand, other fields may contain geographical metadata. For example, the account location and user description are free-form character fields provided by a user that may or may not contain metadata that can be geo-referenced. In order to increase the number of tweets with spatial reference, we can extract location information from the Twitter user profiles, which may represent where that particular user might be active most of the time. Our analytical technique attempted to geolocate the tweets without exact coordinates to leverage these abstract textual fields.

The user location and user description are fields that can be edited by the user and have a maximum length of 30 and 160 characters, respectively; thus, the credibility of information relies on the user. Most users use this field as intended; however, some fill it out in their native language while others use creative ways to share their locations, either by coordinates, fictitious locations, or communicating with emojis. As a result, it creates noise for standardized geolocation approaches. These limitations can be significantly alleviated by extensive preprocessing steps, leading to data loss due to the exclusion of critical words. Nonetheless, a key research issue in this area is establishing how to merge information derived from user-generated data with contextual information.

Generally, these values are not frequently altered and do not necessarily describe the tweet's exact location, but they may represent the user's residence or workplace at least on a city level. Our research considers tweets aggregated on a country level and not the exact location within a city or a country. We still find these data valuable as a proxy in instances where no direct location information for a given tweet was provided. To reduce our dataset size and obtain helpful information, we first filtered it for unique rows and ranked them by frequency. We applied spell checking throughout this field, geolocated them, and finally used the built-in map function of Python and a translator dictionary developed by us to transform all location items and to group similar entries.

6.2.2 Transform coordinates

Besides text, some users choose creative modes to indicate their location, which the approaches mentioned above will not extract. Some provide precise geo-coordinates in various formats in the user_location field. Therefore, the analyzed database contained the three primary forms of a coordinate, which can be composed of

a) degrees (integer), minutes (integer) and seconds (integer or real) (DMS)

b) degrees (integer) and minutes (real number) (MinDec) or

c) degrees only (real number) (DegDec).

We used regex pattern search to extract these values and converted them to DMS.

6.2.3 Transform Emoji flags

Another type of extractable user location is based on emoji flags. To minimize user_location's limited input character field, some specify their home country and additional locations as emoji flags, tiny country pictograms. To extract this meaningful data type, we use the Demoji Python library, which consists of codes from the Unicode Consortium's Emoji code repository. Demoji's findall method produces a dictionary in which the pictogram is converted into a text format (e.g., flag: country name).

6.2.4 Ensuring Consistent Spelling

Although our dataset contains multilingual values, we applied the spell checking package Pyspellchecker over the user location and description field to correct mistyping and increase the subsequent geolocating efficiency. Pyspellchecker supports multiple languages, including English, Spanish, French, Portuguese, German, and Russian. Using the Levenshtein Distance algorithm, Pyspellchecker determines all possible permutations of characters within a predefined processing distance of a word. In other words, it finds all possible strings of characters by inserting, deleting, replacing, and transposing the characters in a given word. Subsequently, it compares each permutation with a dictionary of frequently used words and their frequency distributions and returns the word with the highest frequency as the most likely correct spelling. Pyspellchecker's default editing distance is two characters, which means it finds all possible permutations of the original word that can be created by editing it no more than twice. This is an essential advantage of this algorithm as it corrects mistyping, thus improving the next geolocating step efficiency.

6.2.5 Geoparshing and geocoding with Mordecai

A subsequent step after the spell checking is to extract and standardize possible locations that may represent in the individual Tweets. The approach of geoparsing extracts places from text and matches them to a known place using a global place index. Our approach is based on Mordecai, a full-text geoparsing system that extracts place names from text, resolves them into their correct entries in a gazetteer and returns structured geographic information for the resolved place name. Mordecai was created to provide several features missing in existing geoparsers, including better handling of non-US place names, easy and portable setup and use through a Docker REST architecture, and easy customization with Python swappable named entity recognition systems (NER). The main technical advancements of Mordecai are based on a language-agnostic architecture that uses word2vec (Mikolov et al., 2013) to extract the correct country data for a range of locations in a text in easily changeable named entity recognition models (Halterman, 2017). Furthermore, instead of storing information as rows of columnar data, Elasticsearch stores complex data structures where stored documents are distributed across the cluster and can be accessed immediately from any node. Based on Elasticsearch, distributed document store, this solution provides a remarkable speed-up for geolocating compared to other solutions based on web services (e.g., Nominatim).

Our approach uses spaCy's Named Entity Recognition (NER) for location extraction; then, it uses a custom-built Elasticsearch database for querying geodata for the extracted place names (Honnibal & Montani, 2017). Its gazetteer is based on the GeoNames, a comprehensive library of global place names that contain over 25,000,000 geographical names. It infers the proper gazetteer entry for each placename using neural networks based on Keras and trained on newly annotated language data labeled with Prodigy, an annotation tool powered by machine learning. Although the default probability (country_confidence) is 0.6, we modified it to 0.8 to increase output precision. If any matches over this probability limit, the output includes:

- the name of the most probable matching place
- information about the country in which the place is
- the longitude and latitude coordinates

Although Mordecai extracted all possible locations of each row, we considered the first dictionary as the specific user's location after examining various tweets, possibly referring to the home country of residence. Although acquiring multiple locations is an obvious limitation of this approach, we should note that Twitter only provides a small subset of geolocated tweets, as other studies have reported.

It should also be noted that many of these values are missing because users specify nonexistent locations (e.g., world, moon). However, our inspections over multiple samples of the dataset showed that our mixed offline methodology could extract all the locations present from the text. Nevertheless, since the geoparser treats each retrieved geographical location independently, it may overlook main classifications based on the text's overall context. This restriction may occur when an address contains another existing location entity; for example, "Budapest Hotel, Moscow" appears to be a single location in Russia, but the geoparser may treat it as two distinct entities, 'Budapest Hotel' and 'Moscow.' We did not discover any other issues similar to this one during our random sample inspection of the dataset.

Although, according to the Twitter Developer Platform, approximately 1-2% of Tweets are geo-tagged, while ~30-40% of Tweets contain some profile location information. The above described geoparsher method was able to identify 346,162 places from 655,423 non-empty rows of the user_location field of our Belarusian dataset. Next, we applied the same geoparsing approach to the user_bio field, which returned 10,353 out of the remaining 309,261 rows. The idea behind analyzing this field was that users tend to use it to indicate their occupation and their employer. Overall, our mixed-model geolocating approach was able to identify 356,515 geolocated places which are 54,3% of the Belarusian dataset. Moreover, in our raw Slovakian data set, out of all the tweets, only around 0.2% (24) hold coordinates; however, with the transformation as mentioned earlier approaches, we could provide coordinates for 8069 tweets, which is 61.2% of the original dataset. The remaining 38.8% of the tweets were either posted outside of Europe, or it was impossible to locate them, whereas most of them were excluded as part of the text cleaning process described in Section V.3.1.

6.3 Translation Using Google API

The majority of the tweets were in a language other than English, and as a result, they had to be translated before being used in the subsequent analysis steps. The TextBlob text-processing library, written in Python, was used to translate the tweets that were not in English at the time. TextBlob (Loria, 2018) can be used for various tasks, including part-of-speech tagging, parsing, sentiment analysis, spelling correction, and translation, according to the software's documentation. The API of Google Translate, the most widely used online translation service, is used by the algorithm. It has a pre-trained model that instantly identifies highly accurate languages and can translate them into more than one hundred target languages, including all of the European languages we used in this analysis. This is one of the most significant advantages of using the Translation API. In 2011, as part of a comprehensive accuracy evaluation, 51 languages were translated from one language to another using Google Translate, and the results revealed that the majority of European languages produced reliable results in the majority of cases. With the implementation of a Neural Machine Translation (NMT) model in 2016, the translation accuracy score increased from 3.694 (out of 6) to 4.263, which is very close to a human-level score of 4.636 (Aiken, 2019) thanks to a service update in 2016. This high value is acceptable for the subsequent steps in our workflow because our approach extracts sentiment values from text using a word-based analysis method.

6.4 Emoji/Emoticon Transformation

In general, text pre-processing approaches remove emojis (small images) and emoticons (facial expression representation using keyboard characters and punctuations) from the text. The main problem of such approaches is that users use these small images and characters as the lingua franca of social media to express feelings or ideas, compressing a meaningful word in a short number of characters (Hasyim, 2019). Our methodology does not consider their removal as the appropriate solution since emojis and emoticons contain valuable information, particularly for the subsequent sentiment analysis. Thus, we convert them to text format using the Python emote library to preserve the emoji information for further analysis steps.

6.5 Semantic Analysis

The semantic text analysis process used in our approach is divided into two stages: first, we extend the list of stop words in the algorithm based on the characteristics of our data set, and then we remove these words from the text, and second, we provide a dictionary-based sentiment analysis, which classifies the subjective sentiment information contained in each tweet.

6.5.1 Removing Stop Words

The literature considers auxiliary verbs, conjunctions, and other parts of written text that do not bear the significant semantic meaning as “stop words”. The Natural Language Toolkit (NLTK) (Bird et al., 2009) predefined a list of these words in the algorithm we used. Nonetheless, we added other words to the stop word dictionary that are unique to the unedited text or the analyzed corpus, including unique first names “Martina,” “Marian,” “Andrej,” or words like “gonna,” “wanna.” We remove these words from the dataset and words with three or fewer characters, as they also have limited semantic significance.

6.5.2 Sentiment Analysis

Sentiment scores are used to identify how positive or negative the text of a given tweet is. This identification is performed by calculating the difference between the quantity of positive and negative terms using a vocabulary with positive and negative words automated. We selected the VADER (Valence Aware Dictionary and sEntiment Reasoner) lexicon, a rule-based sentiment analysis tool that uses a lexicon-driven method and heuristics to assess the input data. This method is standardized to the sentiments presented in social media, and it has a higher classification accuracy than other methods in light of the recent literature (Hutto & Gilbert, 2014). For example, in the comparative study of Hutto and Gilbert, it was found that the VADER correlation coefficient (r) was nearly equal to human raters' performance ($r = 0.881$ vs. 0.888). However, when they inspected the classification accuracy for social media text analysis (F1) it outperformed the human raters ($F1 = 0.96$ vs. 0.84) and the other eleven highly regarded analysis tools such as the Hu-Liu04 opinion lexicon or WordNet (Hutto & Gilbert, 2014).

Generally, if the score is lower than zero, the sentence or text element is assumed to include a "negative sentiment," whereas it is considered a "positive sentiment" if this sentiment value is higher than zero. If the score equals zero, the sentence is identified as "neutral." The main problem of this arrangement with only three classes is that the algorithm has shortcomings in determining unambiguous negative or positive scores for sentiment rates around zero because they are either indeed neutral or misclassified with a relatively high probability. The VADER algorithm uses the so-called compound score, which is calculated by adding the valence scores of all words in the lexicon, adjusting them according to the given rules of the algorithm, and then normalizing the value to fall between −1 (the most extreme negative value) and +1 (the most extreme positive value). It is a valuable metric if we seek a single unidimensional assessment of a sentence's emotion. Thus, we categorize the tweets with the compound score into five categories:

- Very positive '5'—(0.55–1.00)
- Positive '4'—(0.10–0.55)
- Neutral '3'—(−0.10–0.10)
- Negative '2'—(−0.55–0.10)
- Very negative '1'—(−1.00–0.55)

This does not propose that all the tweets with the sentiment value of 0.1 and −0.1 are neutral; however, we could reduce the number of neutral tweets using this categorization.

6.6 Spatiotemporal Data Processing and Clustering

We also performed spatiotemporal analysis using the tweets to understand the escalation of protests based on social media activity. Most of the tweets in our query results were posted from European countries; therefore, to keep the analysis concise, we only considered countries from Europe, including Russia and Turkey. We kept the data aggregated at the country level, as language and the political characteristics of a country might influence the tweeting behavior stronger than other characteristics at the city level or other finer spatial scales. Furthermore, most of the time (except in Slovakia and Belarus), the protests were held only in the capital.

Once the tweets were pre-processed and filtered (see Section V.3 and Section V.4), we performed clustering to find countries with similar tweeting trends about Kuciak's murder, which would probably also indicate when protests took place or the presence of other influencing parameters such the media or politics. For this purpose, we used Time Series Clustering in ArcGIS Pro 2.8 (Redlands, C. E. S. R. I., 2.8), where time-series data can be clustered based on three criteria: having similar values across time, tending to increase and decrease at the same time, or having similar repeating patterns. By identifying countries with a similar pattern, we might be able to reveal the influencing parameters and how these parameters changed over time. Moreover, it provides a more concise and informative visualization and interpretation of the result than statistical values for 39 countries one by one.

For our analysis, we performed clustering based on the number of tweets over time for each country to track the tweeting activity of the citizens in general. This means that we considered the second type of clusters (where values tend to increase and decrease simultaneously, but their absolute value is less relevant). For example, a time series with values (1, 0, 1, 0, 1) is more similar to a time series with values (10, 0, 10, 0, 10) than it is to a time series with values (1, 1, 1, 1, 1) because the values increase and decrease at the same time and stay in a consistent proportion. Therefore, we are able to avoid problems related to different population sizes, and no normalization based on the population is needed for the clustering.

6.7 Topic Modeling Using Latent Dirichlet Allocation (LDA) Method

6.7.1 Preparing Steps for Topic Modeling

The first step of topic modeling is tokenization, which is a technique for segmenting texts into smaller units. This algorithm divides the text at each space character to generate a list of separate tokens (unique words, numbers, and signs). We used the Gensim library's simple pre-process function for this step, which iteratively converts tokens to Unicode strings, removing accent marks and lowercasing the string (Rehurek & Sojka, 2011). Tokens shorter than three letters are discarded.

To filter out the most common bi- and trigrams (two and three-word expressions) from a stream of sentences, we used the Gensim library. However, in order to set a proper filtering threshold, we first manually explored these multi-word expressions with the help of the scikit-learn CountVectorizer (Pedregosa et al., 2011), which converts the text to a matrix of token counts. Then, we set up the Gensim threshold to ignore those bi- and trigrams that bear well-known information such as the fact of the murder or the victim's occupation, the location of the assassination, or even the Belarusan events (e.g., "journalist jan," "murder of," "in slovakia," "the murder of", "of journalist jan," „flower revolution") for each cluster identified in the spatiotemporal analysis.

6.7.2 Lemmatization and Vectorization

This step aims to reduce inflectional and derivationally related word forms to a common base form, similarly to a stemming approach. However, in the case of lemmatization, the part of speech of a word (POS tag) such as symbols, numbers, or verbs should be first determined, and the normalization rules will be different for the different parts of speech; thus, it is lexically more sophisticated. This method also involves the grouping of the inflected forms of each word, identified by the word's lemma or dictionary form (for instance, "better" is lemmatized as "good," "cars" as "car"), so they can be analyzed as a single item, thus enhancing the significance of the topic–word associations (Honnibal & Montani, 2017). For lemmatizing, we use the spaCy Lemmatizer (Honnibal & Montani, 2017) that provides a rule-based lemmatization with the setting to allow only proper nouns, verbs, and nouns related to our LDA corpus because our research is concentrating on topics that primarily answer the question of who did what and where. It shall be noted that in an earlier step (see Section V.3.1), we already removed stop words (e.g., auxiliary verbs), which increased the speed and accuracy of the sentiment analysis; however, this step is not necessary for topic modeling since the spaCy Lemmatizer is effectively capable of filtering out certain POS tags, such as auxiliary verbs.

As a final step, the text corpus has to be converted into a vector format because LDA requires a document-word-count matrix and a word dictionary to create a "bag-of-words" corpus, i.e., a collection of words without information on the proper syntax.

6.7.3 Performing the LDA Topic Modeling

We used LDA with the Gensim library for topic modeling on the final set of geolocated tweets in Python (Campbell et al., 2015). There is no generally established a priori parameter modeling approach for LDA. LDA supposes that each document d of a set of documents D includes one or more topics z, characterized by a probability distribution of single words w, the only observed variable in the model. Consequently, the latent variable ϕ describes a multinomial distribution of words within a topic. The other latent variable θ describes a multinomial distribution of topics in a document. α and β are two concentration parameters – α describes prior knowledge regarding the distribution of topics in a document, whereas β holds prior knowledge about the distribution of words in a topic. A higher value of α shows a more plane distribution of topics over a document, while a lower value, significantly lower than zero, shows a higher concentration of topics (Griffiths & Steyvers, 2004).

In order to find the most suitable parameters, the alpha, beta, and the number of topics extractable from the dataset, we apply hyperparameter optimization that seeks the best setting in a validation corpus set (75%). We used the topic coherence measure (C_v) for performance comparison, which is considered to have the strongest correlations with human ratings (Röder et al., 2015). The value of C_v combines an indirect confirmation measure that uses normalized pointwise mutual information (NPMI), cosine similarity, a

Boolean sliding window, and the one-set segmentation of the top words. We have applied this optimization approach to all clusters identified in the spatiotemporal analysis. Our hyperparameter optimization returned the following parameter settings:

- All cluster: α = symmetric, β = 0.91, and number_of_topics = 8.
- Cluster 1 = α = 0.31, β = symmetric, and number_of_topics = 8.
- Cluster 5 = α = 0.01, β = 0.91, and number_of_topics = 7.

The similar hyperparameter settings indicate that the corpus generally follows symmetric distribution in which the lower alpha indicates fewer topics while the high beta represents increased topic-word density; consequently, the discussion revolved around a few themes. Finally, the tweets were classified according to the topic that produced the highest probability; then, we generated the 10 most frequent keywords for each topic. Keywords, however, sometimes are not able to make proper sense of the discussed topic, to overcome this limitation, we also assigned the most representative tweet assigned to each topic.

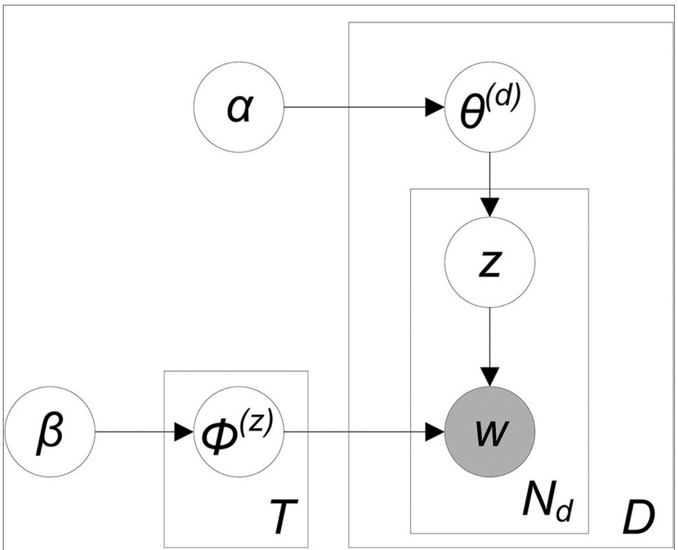


Figure 2: Plate notation of Latent Dirichlet Allocation (LDA) process (based on Griffiths & Steyvers, 2004)

6.7.4 Another topic modeling option with Bertopic

Although Latent Dirichlet allocation (LDA) and Non-Negative Matrix Factorization (NMF) are the most widespread topic modeling techniques, nowadays, we tried another method, which was forced by our second larger Belarus dataset. LDA uses a probabilistic approach, whereas NMF utilizes the matrix factorization technique; however, new techniques established using Bidirectional Encoder Representations from Transformers (BERT) also exist. In this dissertation, we also tried the open-source BERTopic model (Grootendorst & Reimers, 2021) with the pre-trained language "xlm-r-100langs-bert-base-nli-stsb-mean-tokens" embedding model (Reimers & Gurevych, 2019) for topic modeling. XLM-R was trained using the masked language modeling (MLM) objective on an unlabeled text from CommonCrawl datasets in 100 languages, which makes it fit our multilingual social media dataset.

The BERTopic topic modeling approach contains three consecutive stages: extracting document embeddings, clustering, and topic representation. The embedding model extracts the contextualized word representation for all tokens and then passes it to BERTopic, which uses the Uniform Manifold Approximation and Projection (UMAP) algorithm to reduce the embedding dimensionality. Furthermore, it preserves both the local and global structure of embeddings. As a subsequent step, the density-based algorithm (HDBSCAN) clustered the tweets, allowing the identification of the outliers. Finally, the class-based TF-IDF starts looking at TF-IDF from a class-based point of view instead of individual documents, thus supplying all documents within a single class with the same class vector. Practically, class-based TF-IDF is a TF-IDF formula adopted for multiple classes by joining all documents per class. Then, the frequency of words t is extracted for each class i and divided by the total number of words w. Next, the total unjoined number of documents across all classes m is divided by the total sum of the word i across all classes.

$$c - TF - IDF_i = \frac{t_i}{w_i} \times \log \frac{m}{\sum_j^n t_j}$$

Figure 3: Class-based term frequency-inverse document frequency (TF-IDF, based on Grootendorst 2020)

Seemingly this method has advantages over the popular approaches mentioned above. First, there is no need to define the number of topics in advance as class-based TF-IDF extracts the number of topics described in the documents. In addition, Abuzayed and Al-Khalifa recently compared different topic modeling techniques. They evaluated the results of topic modeling techniques using the Normalized Pointwise Mutual Information (NPMI) measure and found that Bertopic showed better results than NMF and LDA (Abuzayed & Al-Khalifa, 2021). However, Bertopic could not catch the hidden topics in the Belarus dataset, so we still use LDA over that database.

RESULTS

7 #Allforjan dataset

7.1 General Spatial and Temporal Characteristics of the Tweets (RQ1)

After performing the pre-processing steps, we had a dataset of over 8000 tweets. **Figure 3** shows how tweet counts varied daily in our analysis period for European countries. The most significant peak was observable on 28 February when Kuciak's unfinished work was published about the connections of the Italian mafia and Slovakian politicians, whereas there are several smaller peaks from 9 March onwards. These most likely represent the main protest day (9 March) and the news around the resignation of the Minister of Interior and the PM in Slovakia (12 and 14 March).

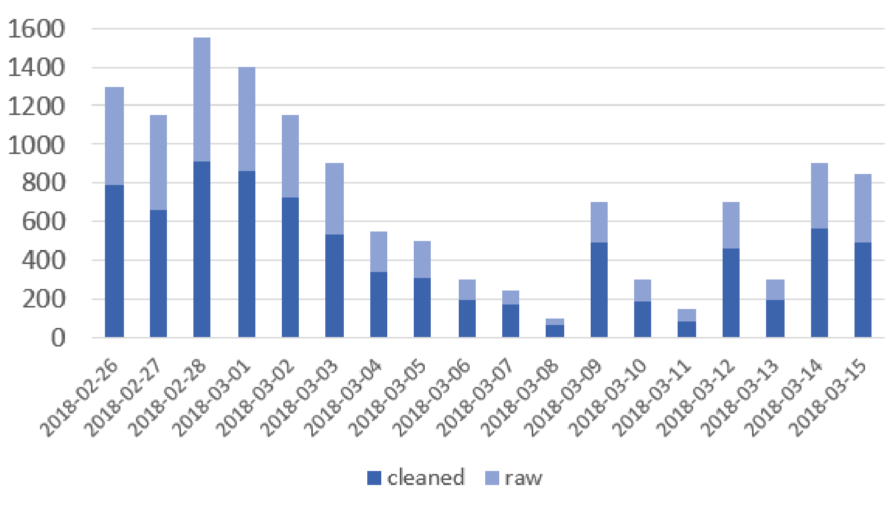


Figure 4: Number of tweets per day before and after the cleaning (#AllforJan dataset)

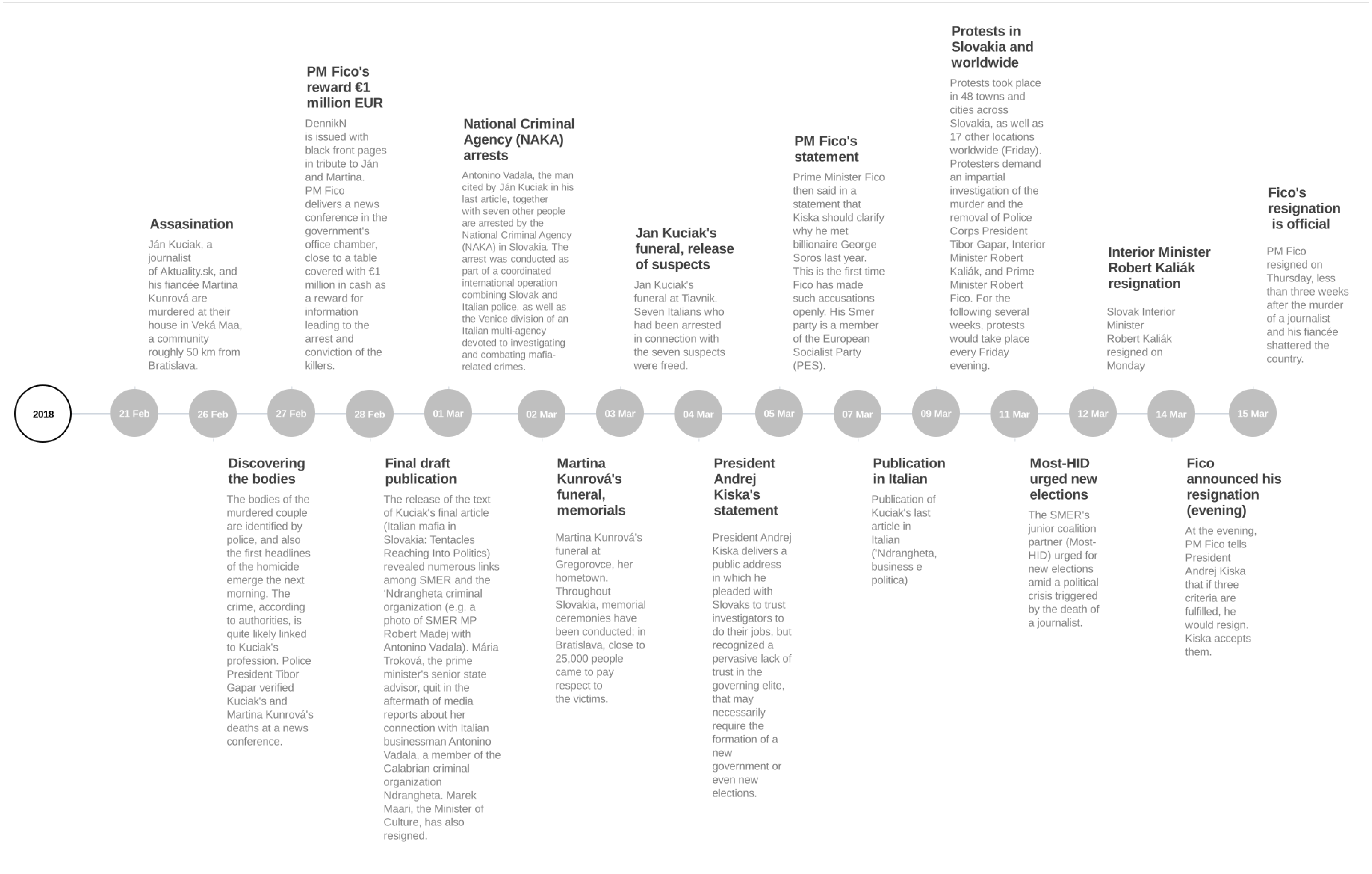


Figure 5: Event timeline (#Allforjan dataset)

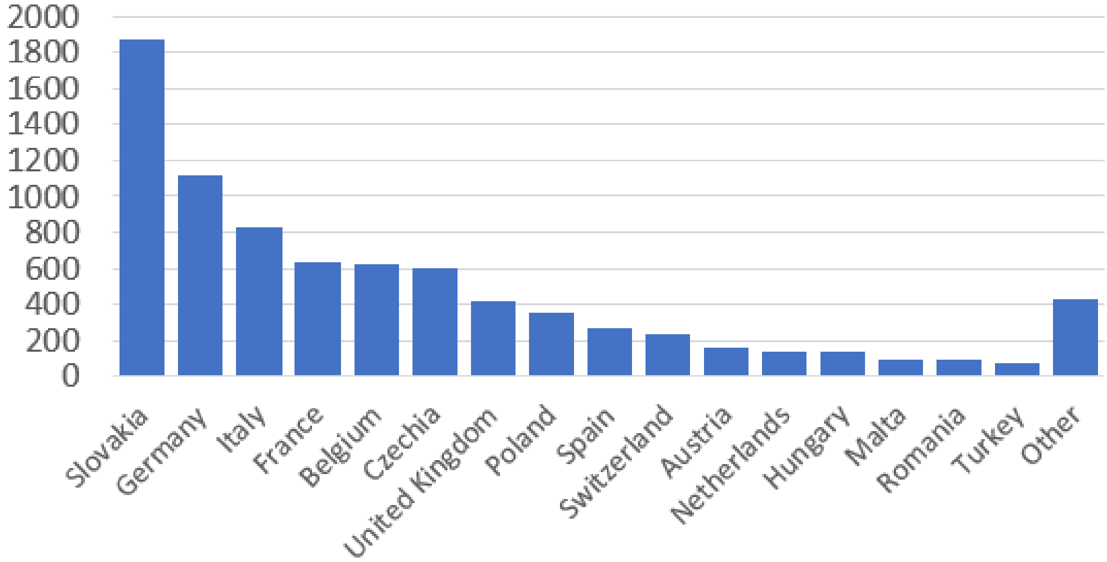


Figure 6: Absolute number of tweets per country (#AllforJan dataset)

Users from Slovakia tweeted over 1800 times throughout the analysis period. The second most active country was Germany (ca. 1100), followed by Italy (around 800 tweets) and France (more than 500 tweets).

Of course, these countries also have large populations, so to exclude the influence of population sizes on the number of tweets, **Figure 7** also visualizes the daily distribution of tweets per country, normalized by their population to make countries more comparable. To calculate the normalized value on each day for a country we used the math.log1p() function in Python, which gives a reliable value also in the case of larger standard deviation and relatively small values.

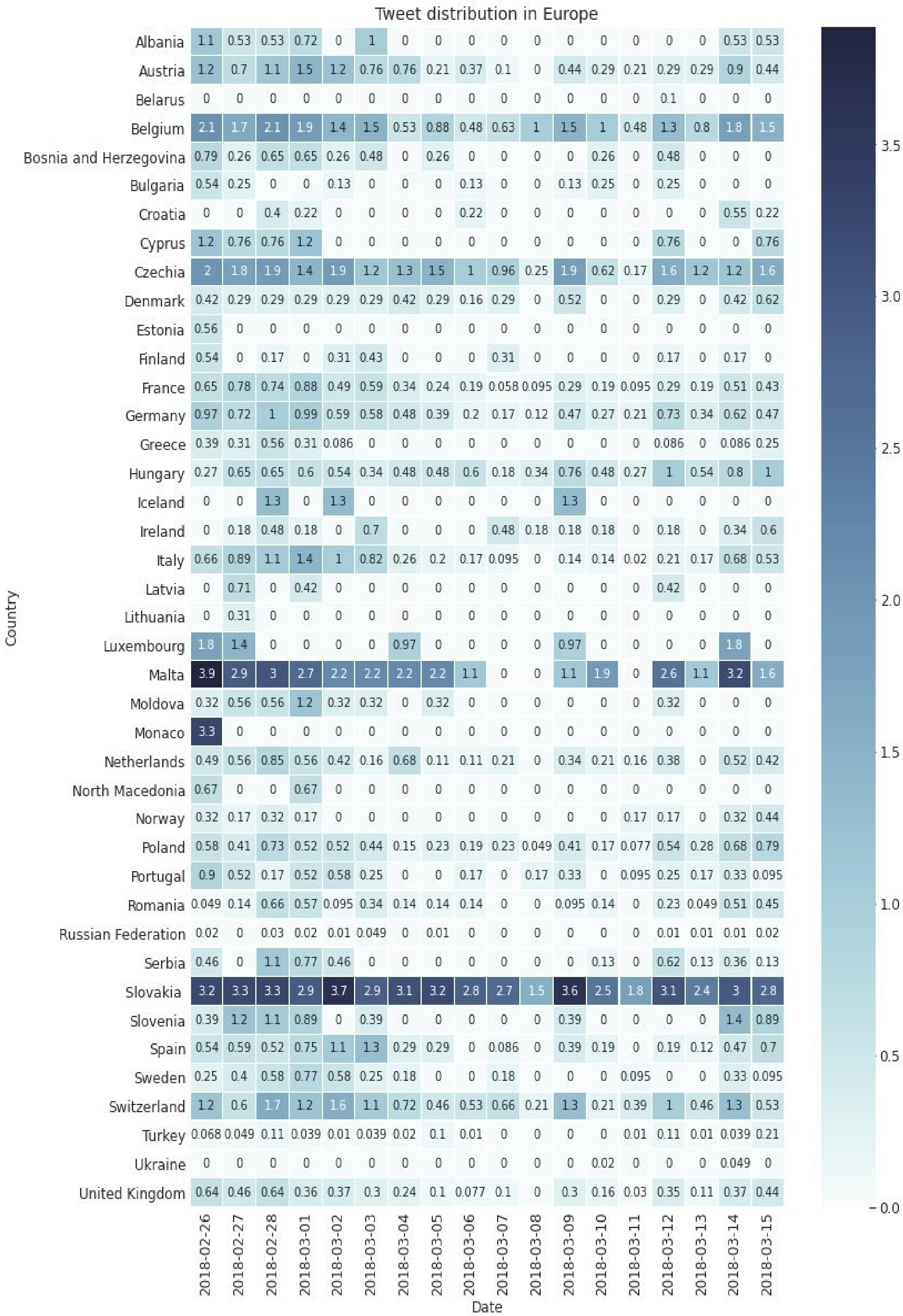


Figure 7: Heatmap of tweets (normalized) per country per day (#AllforJan dataset)

There were eight countries where continuous tweeting activity was observable throughout the 18 days considered for our analysis: Belgium, Czechia, Germany, France, Hungary, Poland, Slovakia, and Switzerland. Czechia, Hungary, and Poland are neighboring countries with shared history in the past, so their interest in the topic can be easily explained. As another neighboring country, users from Austria also had high activity, except for the day of 8 March, when there was no related tweet posted. Germany, the United Kingdom, and France are big countries, and along with Belgium, they represent solid political power in Europe or for the European Union, so this may also explain why they were also actively discussing the case. Although being a small country and far away from Slovakia, Malta also showed high interest in the topic, as a few months prior to the murder of Kuciak a journalist from Malta was also killed because of the investigations she was working on. The role of the Italian mafia was heavily discussed throughout the period due to the corruption among Slovakian politicians, so higher tweeting activity in Italy is also not surprising. The publisher of the journal that Kuciak was working for is in Switzerland, so this probably also explains the high interest there, most likely thanks to the media and news reports. Further statistics about the tweeting activity of users per country can be found in the **Supplementary Materials.**

7.2 Spatiotemporal Clustering (RQ1)

Figure 8 shows the clustering results based on the number of tweets per country and their dynamics over time. Whereas **Table 6** summarizes which countries belong to each cluster and the main characteristics of the time series. The first cluster (blue) contained countries where there was a high tweeting activity (peak: 1 March) in the first few days and a second, smaller peak at the end when PM Fico resigned (14 March). Italy is the country with the most tweets in this group, which is probably thanks to discussions about the responsibility of the Italian mafia in the murder and the corruption in Slovakia.

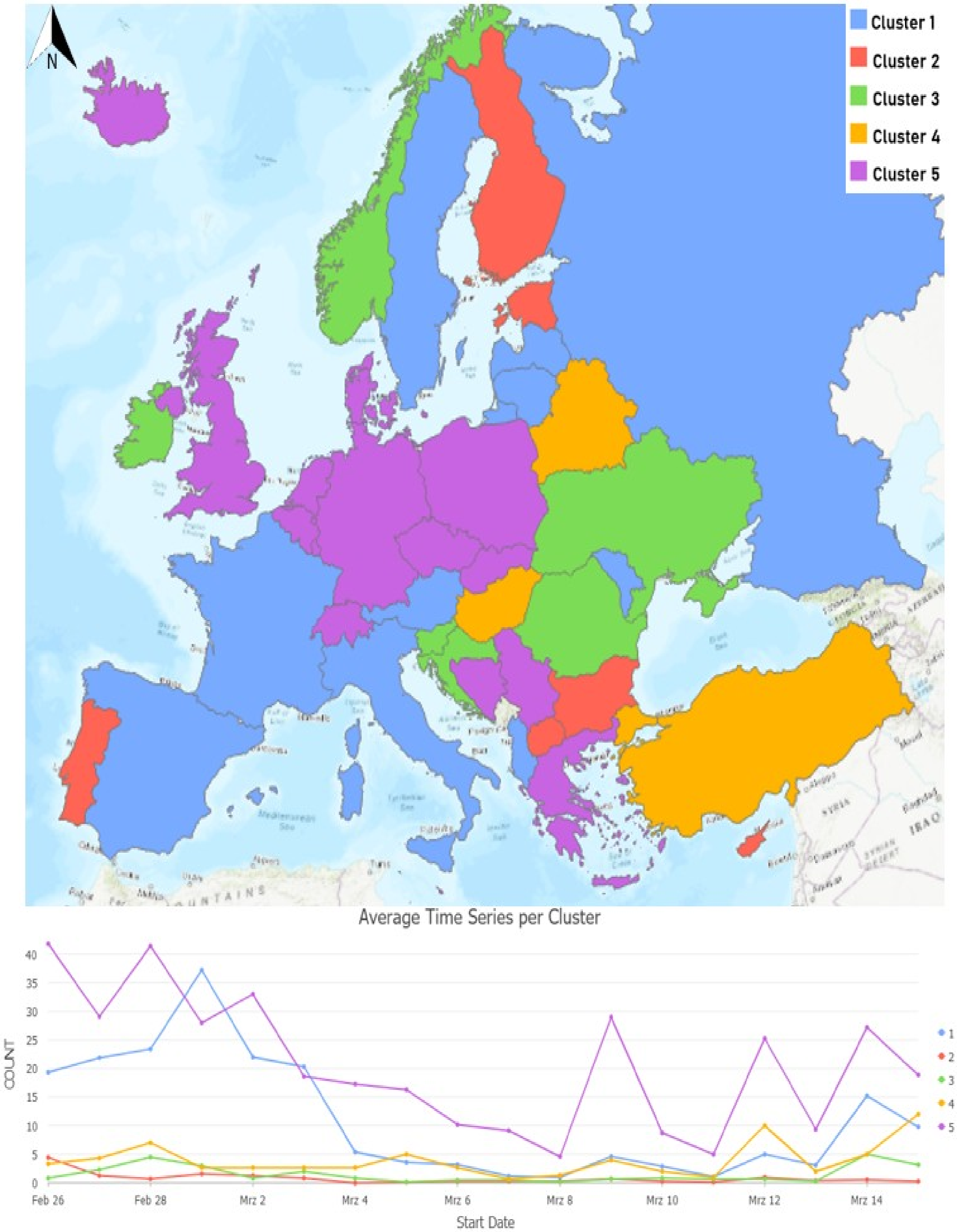


Figure 8: Clusters of spatiotemporal tweeting behavior (#Allforjan dataset)

Users from the countries belonging to Cluster 2 (red) were the least active tweeting related to Kuciak's murder. Tweeting activity in these countries remains low during the whole analysis period. Cluster 3 (green) has a similarly low activity level as Cluster 2, with two smaller peaks on 28 February and 14 March, the two most significant events related to the murder (see **Figure 5: Event timeline**). Cluster 4 has only three countries, and interestingly, there is no peak in the beginning when the murder and its motive were discovered. In comparison, 12 and 15 March are the peaks for these countries that are more related to the resignation of the Prime Minister and other indirect influences of the murder or the journalist's work. If we look at the topic modeling results of these countries, we can observe that there is a high chance that tweets discussing the murder and the following events might have a solid political narrative rooted in the political systems of the countries in this cluster. Although the trend is clear, the absolute number of tweets is relatively low (similarly to Cluster 3), which might also show that Twitter is not the most popular social media platform in these countries. Therefore, those who do use it, might be even less representative to the general population than in other countries, where significantly more tweets were harvested, potentially leading to a distortion of the results. Thus, in the more detailed analysis of the sentiment patterns and topic modeling, we exclude these three clusters (Cluster 2–4) because they do not have enough tweets for such an in-depth analysis. On the contrary, Cluster 5 (purple) has the most tweets (over 5500) and is the group Slovakia also belongs to. In this group, the highest peak was in the beginning, when the journalist was found dead and his unfinished work was discovered and published, then the activity slowly started to decrease, with the lowest number of tweets on 8 March, after which it started to increase again and reach some secondary peaks on 9, 12, and 14 of March.

Table 6: Details of the clusters (#AllforJan dataset)
countries with the most tweets in each cluster are written in bold

Cluster	Countries	Peak (avg. count per cluster	Number of Tweets	Summary
1	Albania, Austria, France, **Italy**, Latvia, Lithuania, Moldova, Russia, Spain, Sweden	1 March and 14 March	2002	First lint to Italian mafia are considered (1 March)
2	Bulgaria, Cyprus, Estonia,	26 February	103	Least active group,

	Finland, Monaco, North Macedonia, **Portugal**			activity remains low throughout the whole period
3	Croatia, Ireland, Norway, **Romania**, Slovenia, Ukraine	28 February and 14 March	162	Activity remains low throughout the whole period
4	**Hungary**, Turkey, Belarus	March 15 (highest), 12 March	213	No peak in the beginning when the murder is discovered
5	Belgium, Bosnia and Herzegovina, Czechia, Denmark, Germany, Greece, Iceland, Luxembourgh, Malta, Netherlands, Poland, Serbia, Switzerland, **Slovakia**, United Kingdom	26 February, 9 March	5591	Larges group, peak in the beginning, slowly decreasing, second half of the period also active

7.3 Temporal Patterns of the Sentiment Values per Cluster (RQ2 and RQ3)

Figure 9 represents each cluster's daily mean sentiment values calculated in Methodology**.** The second half of the period (from March 6 on) is more favorable than the beginning. As Clusters 2–4 have only a few hundred tweets, we mainly focus on Clusters 1 and 5 in detail (**Figure 10**). Overall, Cluster 5 tends to be even more optimistic than Cluster 1.

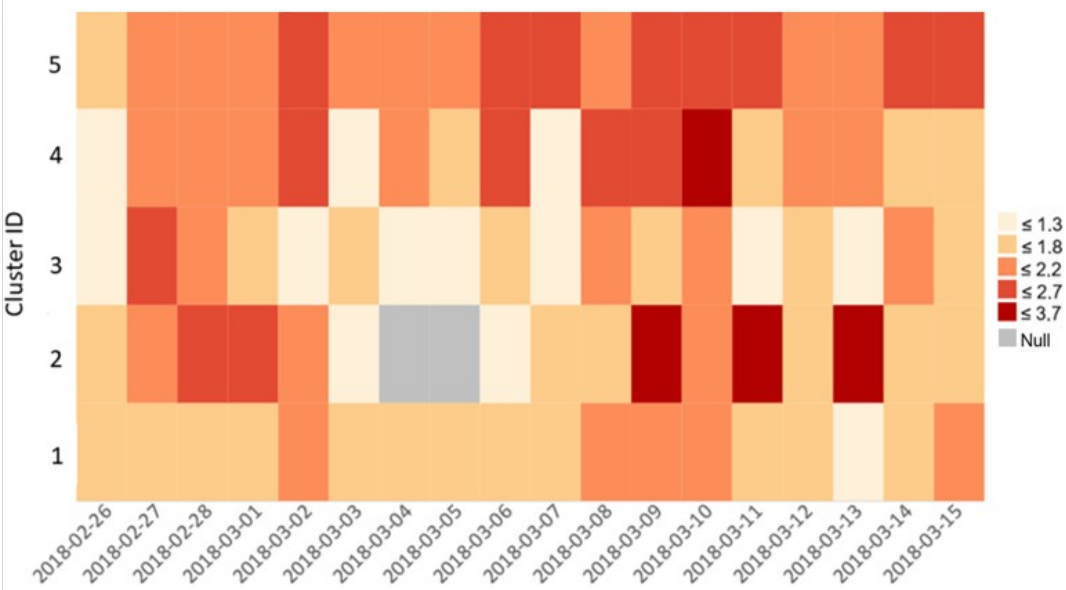


Figure 9: Sentiment values per day for each cluster (#AllforJan dataset)

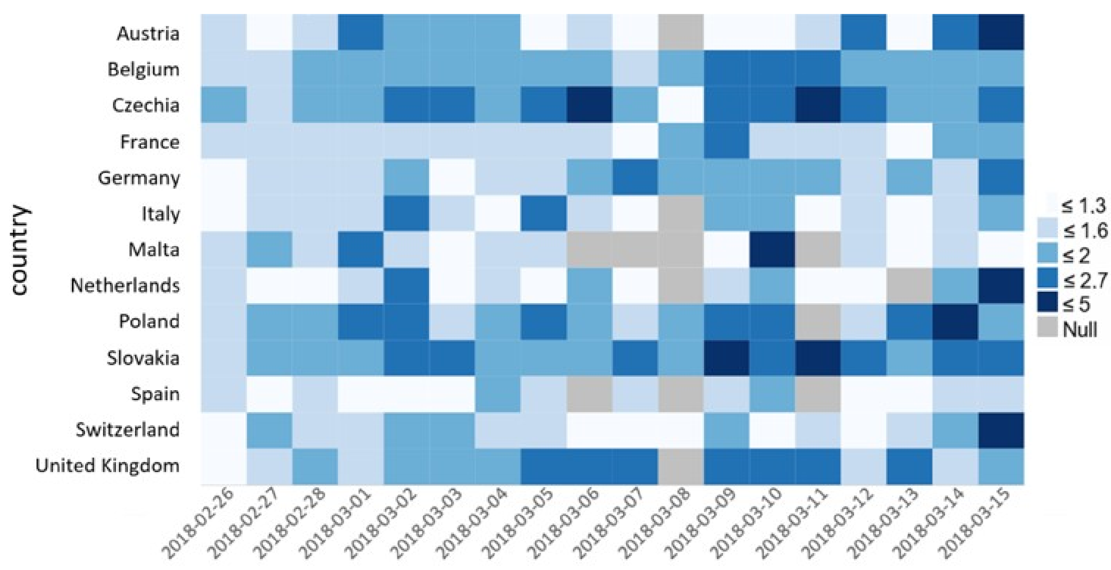


Figure 10: Sentiment class values per day for countries in Cluster 1 and 5 (#Allforjan dataset)

(netural tweets excluded)

Figure 10 shows the country-specific results if we consider the mean compound score classes discussed in Methodology: Sentiment Analysis section (ranging from 1 to 5, 5 being the most positive) for the countries in Cluster 1 and 5 and also exclude neutral tweets (class 3) to highlight the range of sentiments even more. The most positive category occurs after 6 March, and interestingly Slovakia tends to be more positive in this period than any other country. By checking relevant tweets for these days, such as

> “*On friday, 9th March 2018 at 17:00 we will march again we demand a new and trustworthy government. Fico is over” and “finally, the president of slovakia has accepted the resignation of pm robertfico two weeks after the murder of a journalist, and elections are expected to choose a new government”*

We can conclude that most people supported the claims considering the connections between the government and the mafia and were satisfied with the political consequences, such as the resignation of the PM.

Additionally, if we check statistical significance for these trends using the original calculated compound score (before applying sentiment classes), we found that Slovakia and Germany have this among the countries where there was at least one tweet each day in the analysis period, an increasing trend also statistically verified. (Germany had a 95% confidence level, whereas Slovakia was 99%).

7.4 Result of the Topic Modeling per Cluster (RQ2 and RQ3)

Figure 11 shows the eight most significant topics identified based on the tweets in Cluster 1. The topics touch upon the events and news related to the peaks, such as the resignation of the PM, the Italian mafia, or the role of the European Union. Overall, the most significant topic was identified in tweets that condemn the murder, they are followed by worrying voices about press freedom and security. The third most discussed topic in Europe was Kuciak’s article, which was published after his death on February 28 in English and Slovakian, followed by in other languages later on (e.g., in French). The article revealed several connections between the Slovakian governing elite and organized crime. Remaining topics discuss further findings of Kuciak’s article (i.e., “Ndrangheta mafia in Slovakia, and the European echoes of the event”). **Table 7** shows the contribution of each topic in percentages, which represents how likely it is that the representative tweet of that topic was discussing that topic or included those keywords. Values around 0.7 (70%) means that there is a 30% chance that the tweet discussed other topic than this.

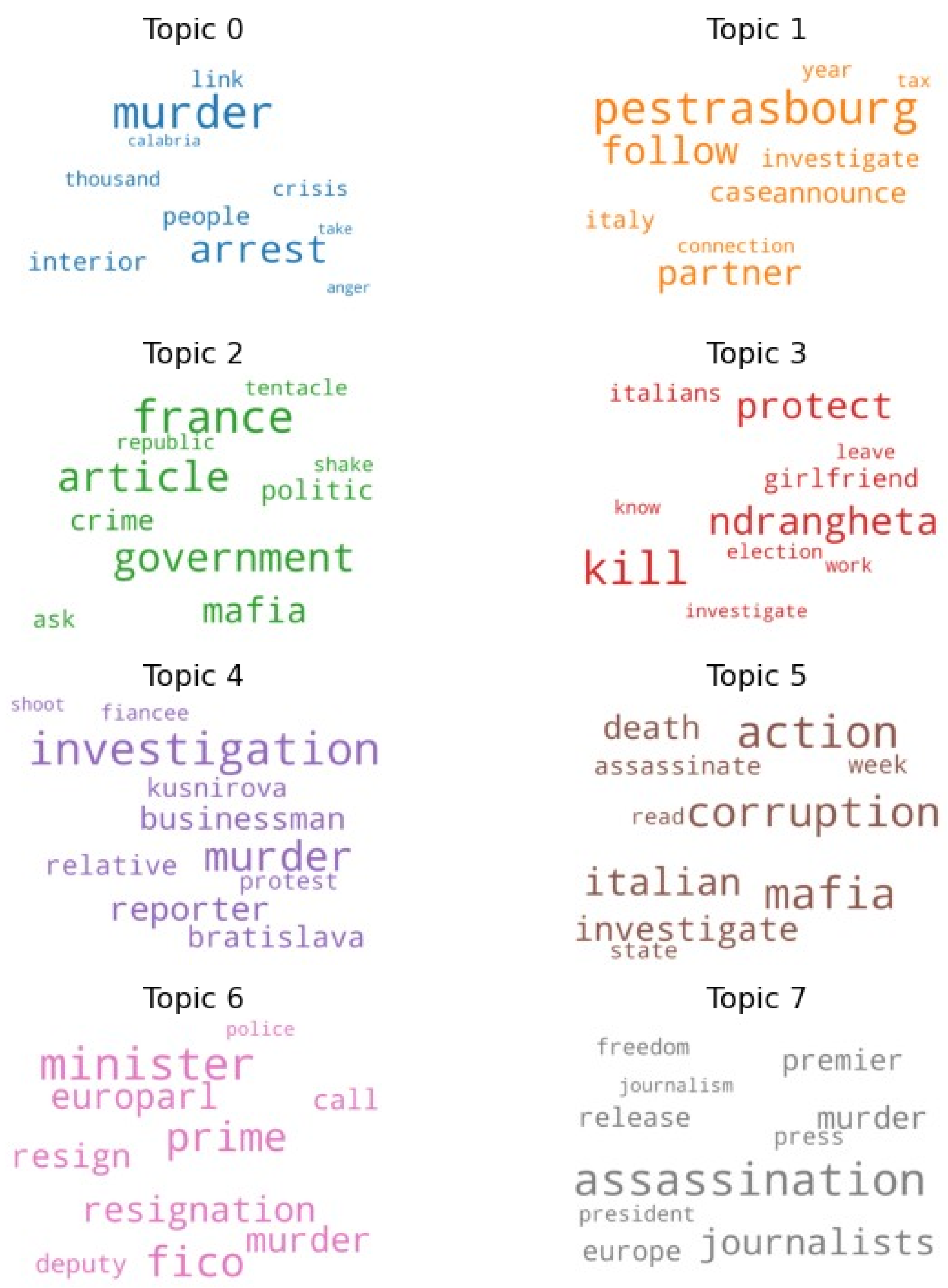


Figure 11: Topic modeling results for tweets in Cluster 1 (#AllforJan dataset).

Table 7: Topic modeling results of Cluster 1 with representative text for each topic (#Allforjan dataset)

(The text is already pre-processed so it might not always make sense grammatically)

Topic #	Topic Contribution%	Keywords	Representative Text
0	0.771081	murder, arrest, interior, people, link, crisis, thousand, calabria, anger, take	in paris a citizens rally takes place on atpm in front of the slovak embassy
1	0.690083	pestrasbourg, follow, partner, announce, case, investigate, italy, year, connection, tax	the slaughter of colleagues continues now its the turn of jan kuciak in slovakia the coordination for the safety of journalists set up in italy is a model in the european states
2	0.770364	france, article, government, mafia, crime, politic, ask, tentacle, republic, shake	the last article of jan kuciak the tentacles of the mafia in
3	0.741201	kill, ndrangheta, protect, girlfriend, italians, election, leave, work, know, investigate	ndrangheta friends of the ndrangheta jan kuciak investigated the infiltration of ndrine in slovakia
4	0.810029	investigation, murder, reporter, businessman, bratislava, relative, kusnirova, protest, fiancee, shoot	we condemn the murder of investigative journalist jan kuciak and his fiancee martina kusnirova
5	0.709867	action, mafia, corruption, italian, investigate, death, assassinate, week, state, read	absolutely sick slovakia is truly going back to its post communists era rip to jan kuciak
6	0.739461	minister, fico, prime, resignation, europarl, murder,	slovak prime minister robert fico resigns after murder of laguardia

		resign, call, deputy, police	journalist
7	0.791084	assassination, journalists, murder, premier, europe, release, press, freedom, president, journalism	president tibor gaspar confirms the murder of jan kuciak might have been related to journalistic work the government offers mil euro as a reward for information about the murder radio slovakia

If we check the topic modeling results for Cluster 5 (**Figure 12**, **Table 8**) there are seven topics identifiable. These topics are more distinct ((**Table 8**) Topic contribution percent values are higher) compared to the topics in Cluster 1, discussing not only the PM but also the whole government's role, the protests, the mafia, the fiancé of Kuciak, and interestingly also Viktor Orban, the Hungarian PM, although Hungary was not in this cluster or among the most active countries in terms of tweeting behavior. The representative tweet of the most significant topic (Topic 2) revolved around the political crisis in Slovakia, especially through the discussion about the resignations and the possibility of new elections. The next topic discusses the situation of press freedom in Europe, and its urgency is further pressed by the fact that some tweets made a direct connection between the assassination of Kuciak and Daphne Caruana Galizia, a Maltese investigative journalist, who was assassinated only a few months earlier, on 16th October 2017 (Rankin & Leroux, 2021).

The overrepresentation of this theme (Topic 4 and Topic 0) shows that the dissatisfied voices about the Caruana Galizia's case further strengthened the Kuciak movement in the online sphere. The tweets that made a connection between the death of Galizia' and the case of Kuciak's fiancée could also be interpreted as a clear representation of the condemnation of the violence against women that may further strengthen this movement. Furthermore, Martina Kusnirova has a double representation among the topics as by name and as fiancée that may suggest that users of this cluster not only strictly condemned her death but they may differentiate between an innocent death and a death related to work. The topic modeling also identified tweets discussing the connection between the Hungarian PM and George Soros, who is a Hungarian-born American billionaire representing a frequent theme in different conspiracy theories and fake news ("George Soros - Wikipedia," 2022). The reason for this relationship may be twofold. First, six months before the assassination of Kuciak, the Hungarian government started a countrywide billboard campaign portraying George Soros and saying “Don’t let George Soros have the last laugh”, thus generating a scapegoat from him regarding the refugee crisis in 2016 (Than, 2017). It may have created a solid base for the Slovakian PM Fico who issued a political statement on 5 March 2018. In this statement the PM inquired into the connection of George Soros and Slovakian President Andrej Kiska, who had declared the possibility of new elections a day earlier. Through this political statement, PM Fico might have tried to discredit the president. The second reason may have been, that PM Orban also saw the “fingerprint” of George Soros behind the Slovakian crisis on 10 March (MTI, 2018). Overall, this representation of Soros among the topics may indicate similar political tendencies among different counties, for example as we have seen that the clustering algorithm put Hungary, Belarus, and Turkey in the same group.

Figure 12: Topic modeling results for tweets in Cluster 5 (#AllforJan dataset).

Table 8: Topic modeling results of Cluster 5 with representative text for each topic (#AllforJan dataset)

(The text is already pre-processed so it might not always make sense grammatically)—

Topic #	Topic Contribution %	Keywords	Representative Text
0	0.999700	journalist, murder, kusnirova, fiancee, investigation, galizia, follow, kill, call, malta	mep chair of ep justice claude moraes takes part in mission to slovakia following horrific murder of jan kuciak amp martina kusnirova the delegation is meeting ngo's journalists amp authorities to understand the case amp look at how eu can better protect journalists in europe
1	0.999572	district, man, sobrance, nothing, orban, set, soros, enter, blame, fakenews	justice and home affairs council jha arrival and doorstep ee urmas reinsalu non cash payments initiative evidence crime
2	0.999739	fico, minister, prime, resign, government, resignation, murder, president, crisis, kalinak	andrej kiska robert fico will remain prime minister even after pellegrini is commissioned to form a new government pellegrini will so far only be the commissioned ie nominated prime minister until the moment before pres kiska appoints his new government then fico will say goodbye to the post of prime minister in resignation
3	0.999647	know, information, police, parliament, murder, question, write, state, accord, request	the head of the anti corruption unit naka robert krajmer was at the scene of the murder of jan kuciak in full police said she corrected her boss tibor gaspar who claimed on Monday that the land surveyor was not in the house or in the

			yard murder of jan kuciak
4	0.999667	freedom, press, democracy, need, europe, journalism, eppgroup, fight, world, let	in the light of recent events democracy remains unimaginable without press freedom stresses publisher detlef prinz jan kuciak daphne caruana galizia
5	0.999700	journalist, murder, mafia, reporter, investigate, arrest, police, article, story, link	jan kuciak was working with occrp reporters looking at a group of businessmen from italy with connections to the ndrangheta mafia a fearsome if successful mafia group our story was never completed due to jan's murder here is part i of why
6	0.999667	protest, bratislava, people, friday, decent, street, kusnirova, demonstration, thousand, slovaks	on friday there will be marches in memory of jan kuciak and martina kusnirova in many cities in slovakia after the announcement of the event in bratislava banska bystrica kosice nitra zilina or krupina and prague were added murder jan kuciak friday marches

8 #Belarusprotest dataset

8.1 General Spatial and Temporal Characteristics of the Tweet (RQ1)

After completing the pre-processing steps, reducing the dataset to European countries, we held a dataset of around 244276 tweets. **Figure 13** shows the primary distribution of tweet numbers among European countries, while **Figure 16 reveals** how tweet counts were altered in our analysis period in Europe. The most significant peak was observable between 15 and 16 August 2020 when Konstantin Shishmakov, the 29-year-old director of Vaukavysk's Bagration Military History Museum, disappeared. As a deputy of the election committee, he declined to accept and sign the falsified election documents; then, he called his wife at 5 p.m and told her that he was going home but never arrived. Shishmakov's body was subsequently discovered in a river in the Grodno region on 18 August. Several smaller peaks are from 15 August onwards (see *Figure 14* for event timeline). These most likely represent the main protest days and relate to events such as the so-called "solidarity chain" protest movements that criticized the post-vote crackdown in Belarus and subsequent police oppression, causing several deaths and detentions.

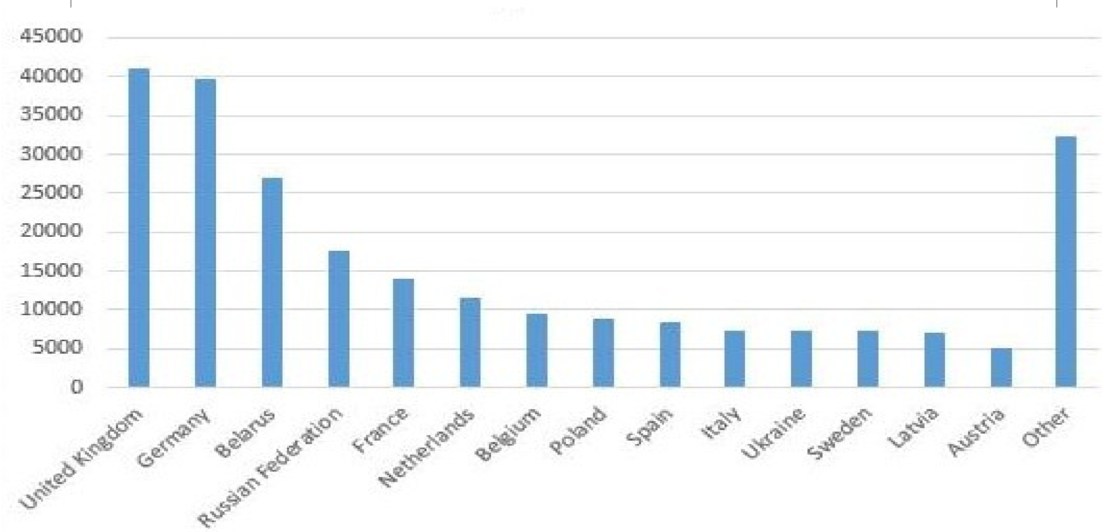


Figure 13: Number of tweets per day (showing the number of tweets in corpus and after data cleaning)

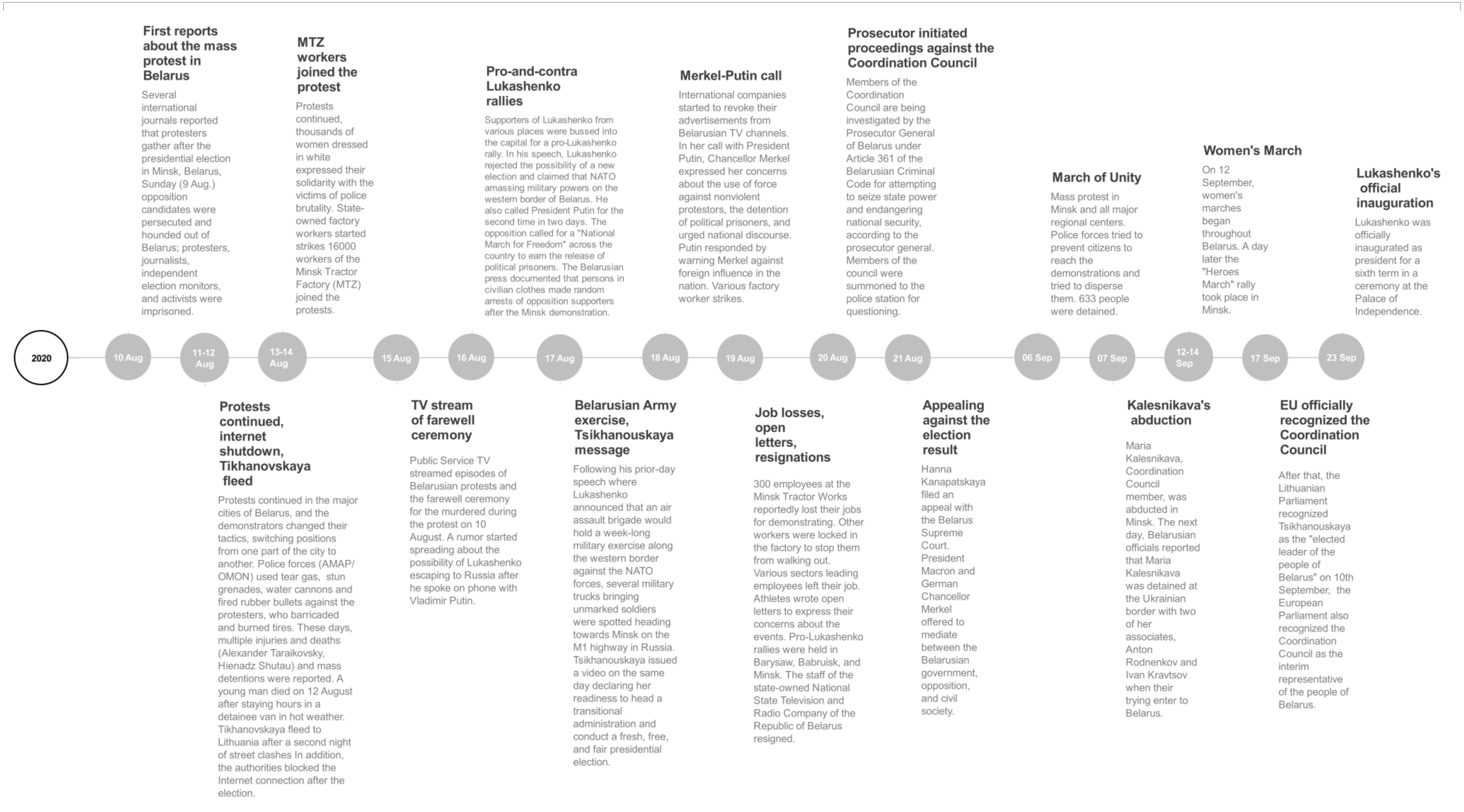


Figure 14: #Belarusprotest event timeline

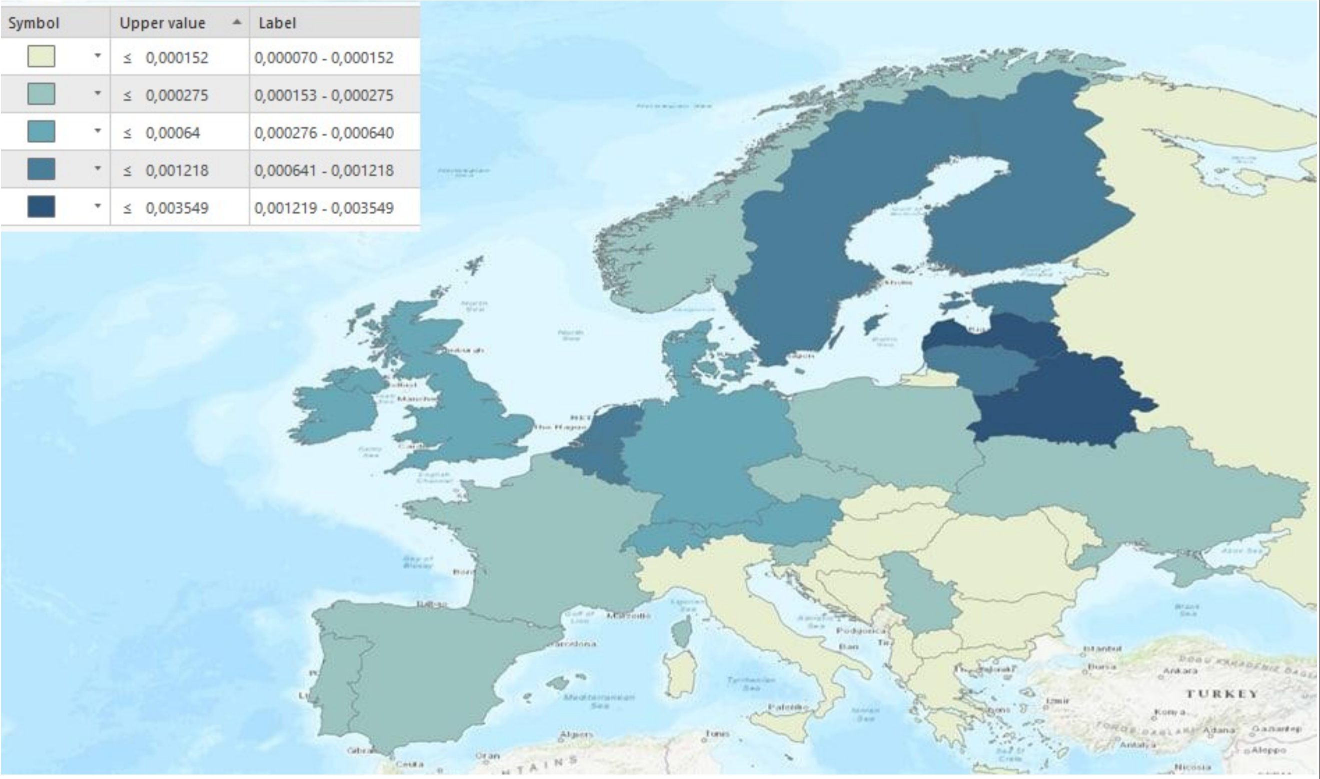


Figure 15: Absolute number of tweets per county (#Belarusprotest dataset)

Users from the United Kingdom tweeted over 40000 times throughout the analysis period. The second most active country was Germany (ca. 39000), followed by Belarus (more than 25000 tweets) and the Russian Federation (more than 15000 tweets).

Naturally, these countries have large populations. To exclude the impact of population sizes on the number of tweets, we normalized them by their population, in order to create a meaningful comparison for the countries. Similarly to our Slovakian dataset, we used the math.log1p() to calculate the normalized value on each day in Python, which gives a reliable value in the case of more significant standard deviation and relatively small values. This result is visualized in the **Figure 16.**

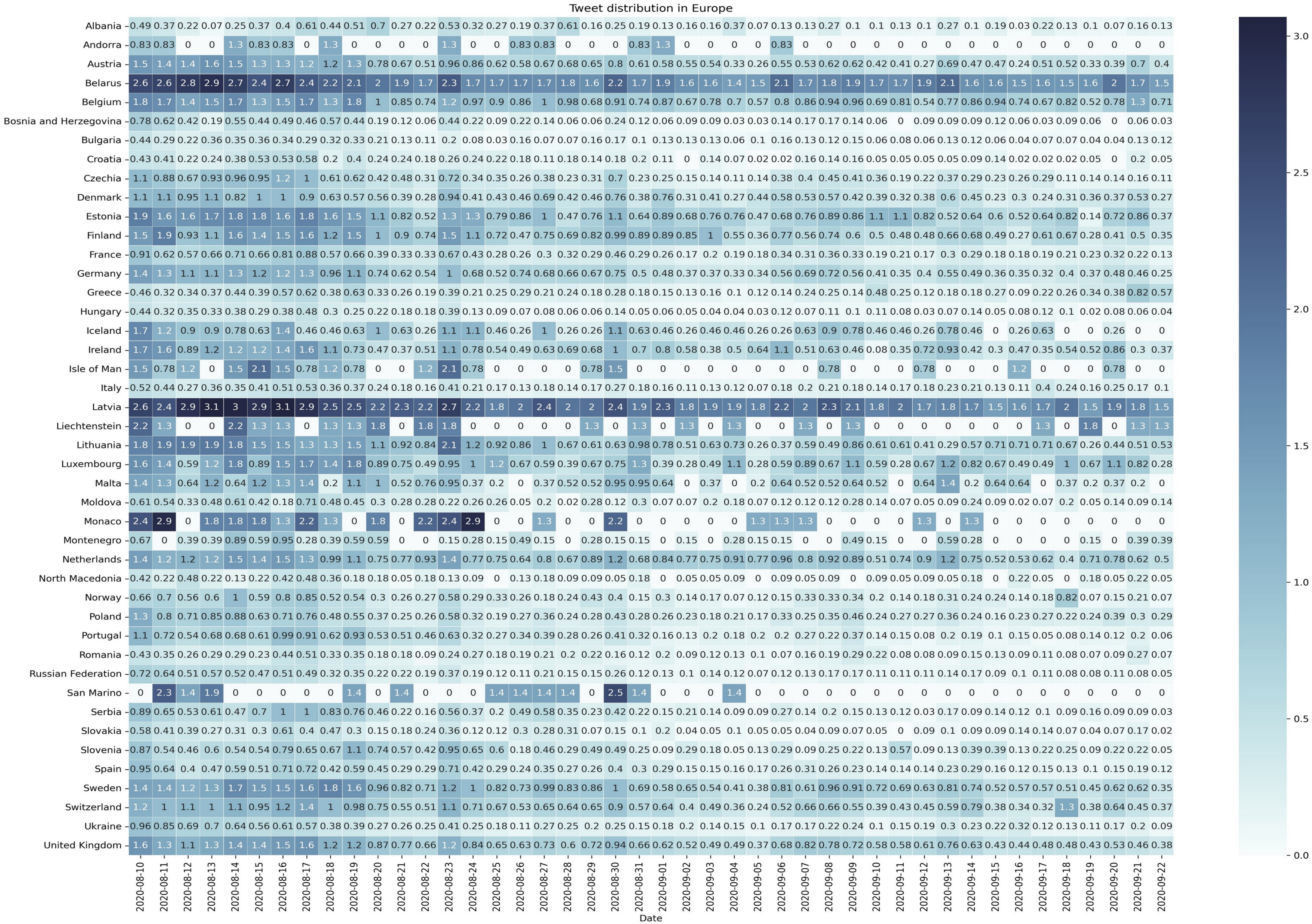


Figure 16: Heatmap of tweets (normalized) per country per day (#Belarusprotest dataset)

There were thirty-three countries where there was continuous tweeting activity observable throughout the 44 days considered for our analysis: Albania, Austria, Belarus, Belgium, Bulgaria, Croatia, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Moldova, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Serbia, Slovenia, Spain, Sweden, Switzerland, Ukraine, and United Kingdom. However, our normalized heatmap visualization emphasizes the most active countries, while on the vertical axis, it makes the most dynamic days among the European countries and other territories visible.

Latvia and Lithuania are neighboring western countries, Estonia is also nearby. These countries were the first who announce sanctions against Belarus. On 18 August, the Lithuanian parliament set economic sanctions, while at the end of the month, Estonia, Latvia, and Lithuania together set sanctions against Alexander Lukashenko and 30 Belarusian officials. These countries incorporated the decision of the European Union. On 14 August in Brussels, Josep Borrell, EU High Representative for Foreign Affairs and Security Policy, informed journalists and a broader audience that the EU would convey sanctions against Belarusian officials responsible for the election falsification and police brutality. On 16 and 23 August mark the main protest days in Belarus, on 6 September, a "March of Unity" took place in the Belarusian capital Minsk with approximately 200,000 participants and all major regional centers. Around 6,000 people took part in the protests in Homel, 4,000 in Hrodna, 3,000 in Brest, Vitebsk, and Mogilev, so this may also explain why they were also actively discussed across Europe. Sviatlana Tsikhanouskaya, who ran during the Belarusian presidential election as the main opposition against Alexander Lukashenko, also fled to Lithuania after publishing the election results. Further statistics about the tweeting activity of users per country can be found in the **Supplementary Materials**.

8.2 Spatiotemporal Clustering (RQ1)

Figure 17 shows the clustering results based on the number of tweets per country and their dynamics over time. Whereas **Table 9** summarizes which countries belong to each cluster and the main characteristics of the time series. The first cluster (blue) contains countries where there was a high tweeting activity in the first seven days; however, after the first peak (18 August), it shows a decreasing pattern with repeatedly lower peaks on 20, 24, 28, and 31 August, as well as 7, 14 and 21 September. Furthermore, there is a more minor but more extended increasing tension pattern between 7 and 9 September. The single peaks are Mondays, reflecting the discussion of protests days held on Sundays.

Moreover, Cluster 1 represents a highly sensible group of political decisions and events. Although there were several mass protests in the first few days, the first turning point was when several state-controlled companies joined the protesters on the streets on 18 August. The day was also a turning point in political communication since the opposition launched the so-called Coordination Council to facilitate a transfer of power. Two days later (20 August), the Council insisted that a new presidential election was needed while the Belarusian chief prosecutor opened a criminal investigation against them. On 24 August, multiple international circumstances affected the online tendencies. First, journals facilitated distributing the information that up to 50,000 Lithuanian protesters attached themselves to the human chain protests a day ago.

Moreover, Tsikhanouskaya met the U.S. deputy secretary while Belarusian police detained two members of the Coordination Council on the same day. 28 August marked the decision of EU foreign ministers to impose sanctions on up to 20 Belarusian officials. On 31 August, Lilia Vlasova, one of the seven members of the Coordination Council Presidium, was detained. Between 7 and 9 September, Cluster 1 users focused on opposition organizers of the previous day's protest, who were kidnapped and forcibly taken to the border with Ukraine in order to leave the country without actually being able to do so.

The significant difference in trend between the 1st and 2nd clusters is primarily detectable in the first part until 24 August. Users from the countries belonging to Cluster 2 (red) were the least active compared to Cluster 1. Tweeting activity in these countries shows a low activity level and decreasing tendency in the analysis period, with three smaller peaks on 24, 31 August, and 7, 14, 21 September. Moreover, its peaks happened a day earlier than Cluster 1 because the users of these countries focused on the protests. On 13 August, the fifth day of the protest, participants held long lines of solidarity.

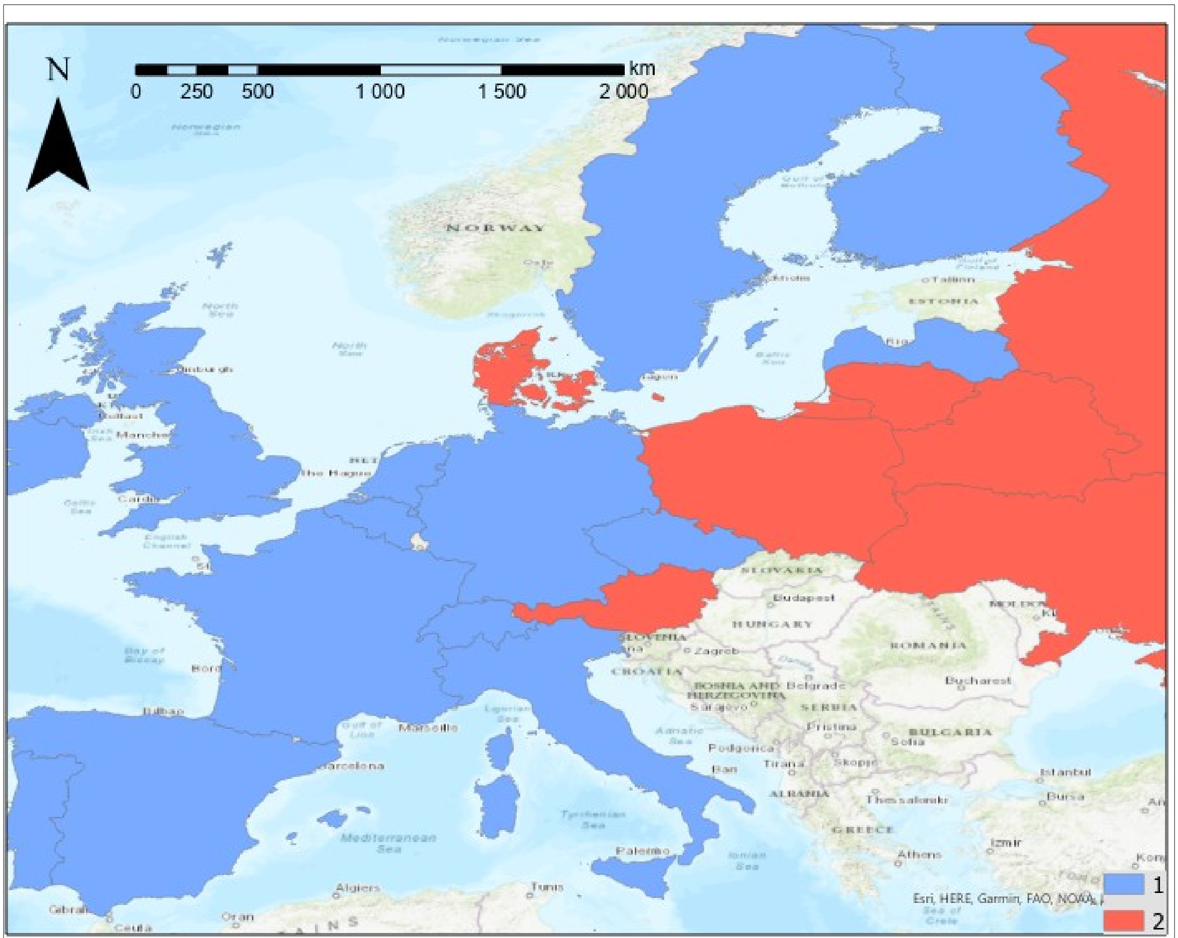
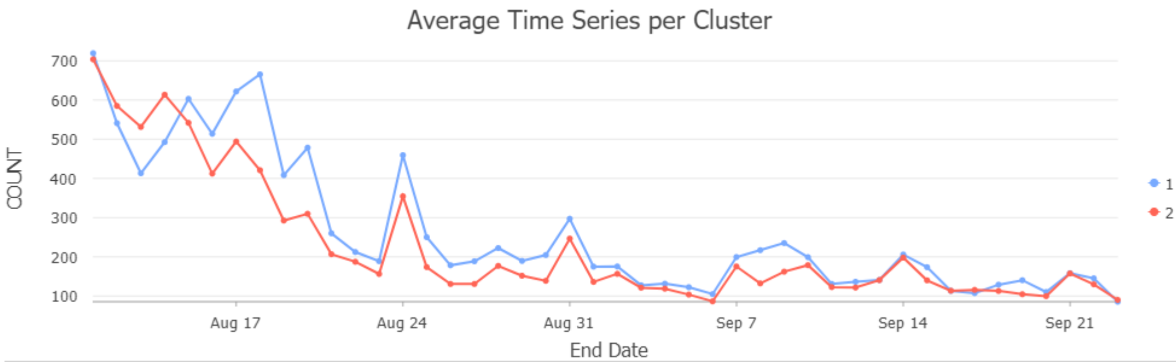


Figure 17: Spatiotemporal tweeting behavior clusters (map and a graph representing countries with above-median tweets over 2000)



If we look at the topic modeling results of the countries of Cluster 2, we can observe a high chance that the tweets discussing the protests, their participant, the insults against them, and the effects of the rally. They might have a solid political narrative rooted in this cluster's countries' political systems. For more details on the interpretation of these topics, see **Section 8.3.2**. Although the trend is clear, the absolute number of tweets is lower than in Cluster 1, which could also show that Twitter is not the most popular social media platform in these countries. Those using it may be even less representative of the general population than in other countries where significantly more tweets were collected, potentially deforming the results. Thus, in the more detailed analysis of the sentiment patterns and topic modeling, we exclude those countries where the median tweet number was under 2000 because they do not have enough tweets for in-depth analysis. On the contrary, we should note that without this exclusion, Cluster 1 has the most tweets (over 162000) and is the group that countries of Western Europe belong to. In this group, the highest peak was happened later, when the European Union inducted sanctions against Belarus, the activity followed the political decisions.

Table 9: Details of the clusters (#Belarusprotest dataset)

(countries with the most tweets in each cluster are written in bold).

Cluster	Countries	Peak (avg. count per cluster	Number of Tweets	Summary
1	Belgium, Czechia, Finland, France, Germany, Ireland, Italy, Latvia, Netherlands, Portugal, Spain, Sweden, Switzerland, **United Kingdom**	18 and 24 August	162051	Belarusian Opposition launced the Coodination Council to facilitate a transfer of power / Tsikhanouskaya met the U.S. deputy secretary
2	Austria, **Belarus**, Denmark, Lithuania, Poland, Russian Federation, Ukraine	14 and 24 August	69868	Authorities released about 2000 people, many of them claimed that they were tortured / EU foreign ministers agreed on the need to place sanctions against Belarus (14 August)

8.3 Temportal Patterns of the Sentiment Values per Cluster (RQ2 and RQ3)

8.3.1 Cluster 1 (Europe)

Figure 18 represents the daily mean sentiment values for Cluster 1 calculated in the methodology sentiment section. The period is slightly negative, with extreme changes between positive and negative values. Furthermore, users of Cluster 1 considered several events more positively in September, especially on 11, 15, 19, and 21 September, which were mainly development stages in the political situation. For instance, the European Union condemned the violence and called on the Belarusian authorities to release all previously detained protesters. On the other hand, the indicators of the most negative values were the most significant protest and their aftermaths in Belarus, such as the long lines of solidarity (13 August) and when three senior opposition organizers went missing a day after ten thousand Belarusian protesters marched through Minsk. This remark is also supported by the visualization of daily and hourly distribution of the tweets which shows a strong concentration on Sundays and Mondays, those when the protests were held.

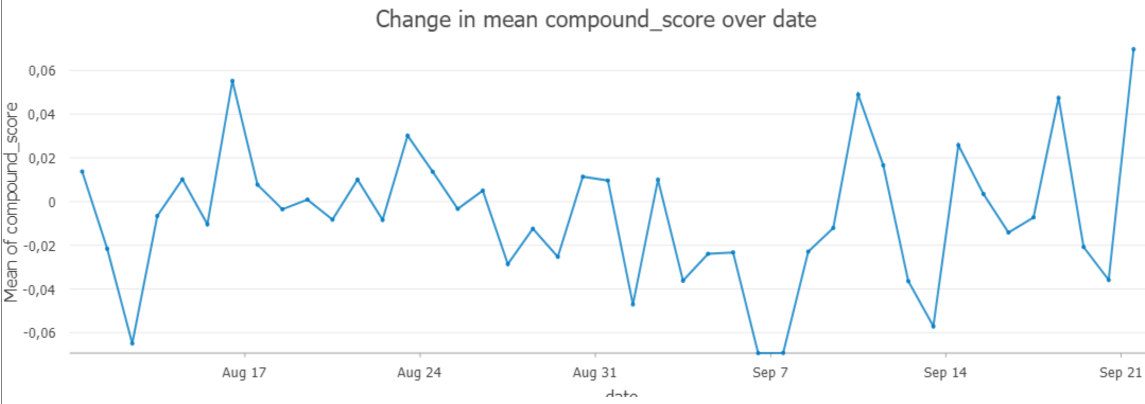


Figure 18: Sentiment values per day for Cluster 1 (West) (#Belarusprotes dataset)

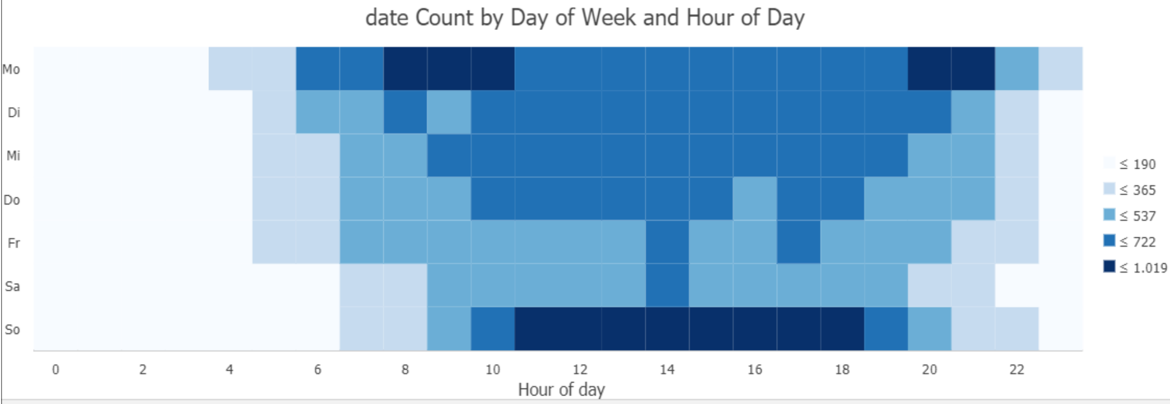


Figure 19: Tweet distribution by day of week and hour of day for Cluster 1 (West) (#Belarusprotests)

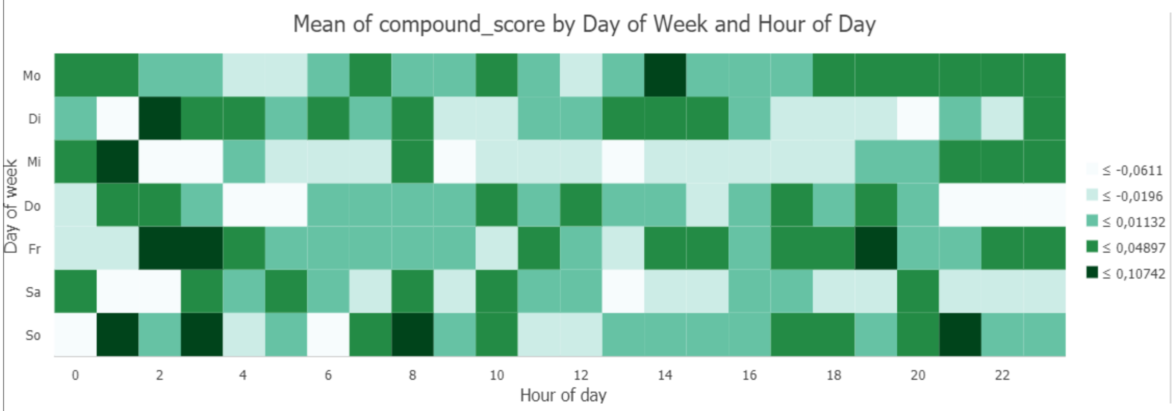


Figure 20: Mean sentiment values per day of week and hour of day for Cluster 1 (neutral tweets excluded) (#Belarusprotests)

Figure 19 shows the daily tweeting activity results, showing that most tweets are concentrated on protest days. **Figure 20** shows the daily specific results if we consider the mean compound score classes discussed in the methodology (ranging from -1 to 1 being the most positive) for the countries in Cluster 1 and also exclude neutral tweets to highlight the range of sentiments even more. The most positive values occur after the early hours and night in the majority on Mondays, reflecting the protests marches. Moreover, users tend to be negative in the intermediate timeframe, especially a day before certain protests, which may reflect their opinion against authorities. By checking relevant tweets for these days, such as

"*Well I was not there, but you learn the lesson and you realize that the world around it even if it is far away affect you, and this is how I repair my mistake. 1 person has already died today in Belarus, thousand are hurt." (negatív)*

„In Belarus, it comes to protests after "Europe's last dictator" Lukashenko wants to win the election victory." (positive)

Considering the analytics data of Cluster 1, we can conclude that most people supported the claims, considering the connections between Lukashenko and the election fraud and the European political response to it. These users were satisfied with the political consequences, such as forming the Coordination Council and the sanctions of the European Union. Additionally, if we check the statistical significance for these trends using the original calculated compound score, we found that among the countries, where there was at least one tweet each day in the analysis period.

8.3.2 Cluster 2 (East)

Figure 7 represents the change in mean sentiment (compound score) for Cluster 2 calculated in Section VI.2.2. The whole period is slightly negative with two extreme negative values on 13 August and 7 September, similar to Cluster 1. The significant difference between Cluster 1 and 2 is that the sentiment values changed more sharply. It is enough to compare the timeframe between 18 and 20 August, which remained settled, similar to the days between 21 and 22 August or the days between 7 8 September. Furthermore, between 15 and 22 September, there is no intermediate peak compared to Cluster 1.

Consequently, the protests strongly affected the users' feelings in Cluster 2, while they were also considered positive when Belarusian factory workers started a general strike. Furthermore, while users of Cluster 1 viewed a positive sign that the European Parliament recognized the Coordination Council as the interim representative of the people of Belarus, users of Cluster 2 focused on the adverse events in Belarus, such as a self-immolation attempt near a police station in Smaliavichy or when 390 women became detained by Belarusian forces on 19 September. Overall, countries and users of Cluster 2 focused on their daily life and suffering in Belarus, while Cluster 1 concentrated on the constant development of the political situation (e.g., sanctions against Belarus; considering the Coordination Council as a legitim power). This remark is also supported by the visualization of the daily and hourly distribution of the tweets that consider mainly Sundays and Mondays as busy days, consequently, those when the protests were held.

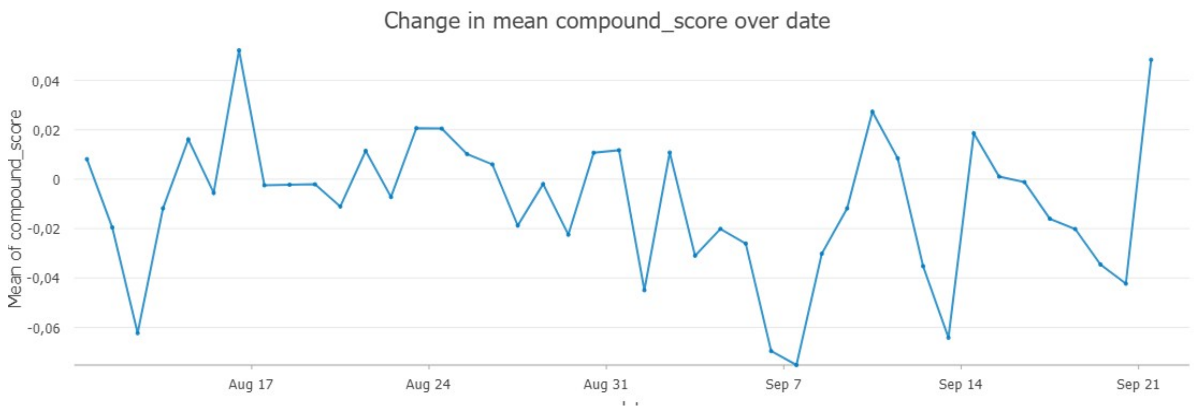


Figure 21: Sentiment values per day for Cluster 2 (East) (#Belaruprotest)

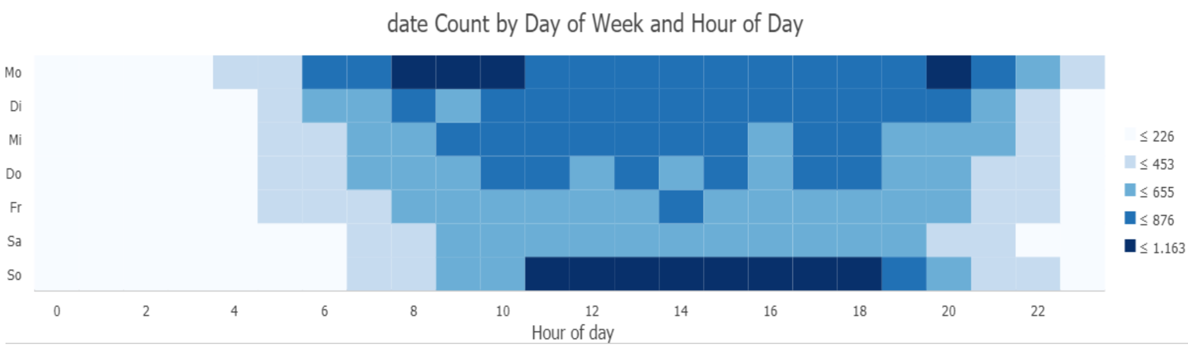


Figure 22: Tweet distribution by day of week and hour of day for Cluster 2 (East) (#Belaruprotest)

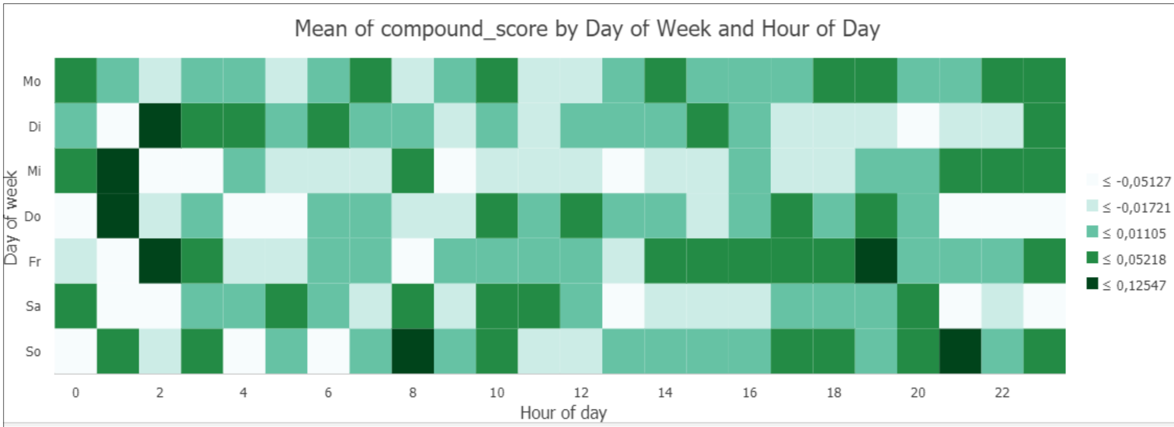


Figure 23: Mean sentiment values per day of week and hour of day for Cluster 2 (East) (#Belarusprotest)

(neutral tweets excluded)

Figure 21 shows the daily specific results if we consider the mean compound score classes discussed in the methodology (ranging from -1 to 1 being the most positive) for the countries in Cluster 2 and also exclude neutral tweets to highlight the range of sentiments even more. The most positive values occur after the early hours and night in the majority on Mondays, reflecting the protest marches similarly to Cluster 1. Moreover, users tend to be negative in the intermediate timeframe, especially a day before a certain protest, which may reflect their opinion against authorities.

"The police van hit a person, seems, he is dead. Minsk (Belarus) now Belarus2020 " (negative)

„Lukashenka falsified the elections again, but Belarusian society in the most majority voted for the opposition candidate, a cichanous.People are accumulated in Belarus to show support for changes.Жыве беларусь!Long live free, democratic Belarus!Belarus" (positive)

We can conclude that most of the people supported the claims considering the different events, protests, strikes, and political decisions from the inside. Overall, Cluster 1 focuses on the wider political decisions and their effect from a European point of view, while Cluster 2 consider the effect of protests and strikes from Belarusian and neighboring point of view.

8.4 Result of the Topic Modeling per Cluster (RQ2 and RQ3)

8.4.1 Cluster 1 (West)

Figure 24 shows the eight most essential topics identified based on the tweets in Cluster 1. The topics touch upon the events and news related to the peaks of this cluster, such as the significant protest events in Minsk, Alexander Lukashenko's connection with Vladimir Putin, or the role of the European Union in the sanctions against Belarus. Overall, the most important topic was identified in tweets that condemned the events in Belarus by leaders of other countries. Such as the president of France, Emanuel Macron, who said Belarus leader Lukashenko "has to go" (Braun, 2020), or U.S. president Biden who slammed Trump for refusing to speak out about Belarus' dictator' Lukashenko (Hunnicutt, 2020). These thoughts are followed by a focus on Lukashenko's cordial connection with the Russian president and its effect on the events and the neighbors of Belarus. It incorporates the topic of when Russia lent Belarus $1.5bn as Lukashenko told Putin, "a friend is in trouble," (Dallison, 2020), which created severe concerns in Lithuania and Poland. The third most discussed topic was EU sanctions against Belarus as well as the effect of the local strike on the economic relations since the EU was the second leading economic partner of Belarus in 2020 (European Comission, 2021). This topic is followed by one that focuses on the need for democracy, especially since Lukashenko is "the last dictator" of Europe (Reuters, 2012). The last topics of this cluster concentrated on the protests, notably when the women held a protest march. **Table 10** shows each topic's contribution percentage, representing the likelihood that that topic's representative tweet discusses that topic or contains those keywords. Values around 0.7 (70%) mean a 30% chance that the tweet addressed a topic other than this one.

Figure 24: Topic modeling results for tweets in Cluster 1 (West) (#Belarusprotest).

Table 10: Topic modeling results of Cluster 1 (West) with representative text for each topic (#Belarusprotest)

(The text is already pre-processed so it might not always make sense grammatically)

Topic #	**Topic Contribution%**	**Keywords**	**Representative Text**
0	0.9635	lukashenko, opposition, president, protest, election, leader, alexander, belarusian, new, vote	an opposition candidate in the disputed belarus election says she ready to become an interim national leader as protests sweep the country in a video message sviatlana tsikhanouskaya said her goal would be to restore calm organise a new vote
1	0.9663	lukashenko, people, go, support, democracy, country, know, freedom, time, want	lithuania, poland and romania firmly support the people of belarus for freedom, human rights and democracy they called the eu to prepare a package of support for the release of the communist dictator aleksandr lukashenko
2	0.9650	minsk, protest, arrest, woman, police, protester, lukashenko, detain, today, people	please pay attention to Belarus Sundays official exitpools show 80 support for a dictator lukashenko who takes away the freedom the language the culture of people living in belarus special forces omon are turned against protesters on the protest day
3	0.9650	covid, lockdown, news, example, pressure, attempt, interview, netherlands, deal, host	covid 19 tdf2020 disneyplus mufc schroeder cancelculture twitch serdarsomuncu couchpeloton prinsjesdag2020 coronavirus appleevent2020 belarus esken fifa21 Indiegame jkrowling lockdown sustainability

			onepiece ps5 lateral thinking radioelemet trumpknew xboxseriesx
4	**0.9786**	de, lukashenko, lukaschenko, la, macron, europe, europa, die, para, biden_trump	in times of rebellion against the dictatorship regime in belarus there is reason to remind that sd's leadership is the same scrap and grain, wants to abolish the independence of authorities officials restrict freedom of the press, force people they do not consider swedish to leave sweden
5	0.9676	eu, sanction, cyprus, turkey, situation, human_rights, say,	because one did not do it in ukraine who in every way is closer and potentially integrable in the eu, and we do not even give the stuff a marshall plan I think they take that risk on belarus gets a marshall plan one can probably blast the eu of pure resentiment.
6	0.9676	number, strike, union, self, white, worker, company, tomorrow,	more and more sources report that belarus regime has no clue what to do about the worker strikes the people joining the revolution peacefully clear is that they lost control rumors are that families of regime are fleeing the country even military are on the side of the people
7	0.9718	russia, putin, lukashenko, russian, ukraine, border, poland, lithuania, china, moscow	lukashenko considers himself exempt from the way we do things in europe he enjoys the protection of the kremlin putin and that is why people in belarus will die trying to stand up for their rights the world is watching lukashenko step down now

8.4.2 Cluster 2 (East)

If we check the topic modeling results for Cluster 2 (**Figure 25**,**Table 11**), eight topics are identifiable. These topics are distinct (**Table 11**) The percentage values of the topic contributions are higher) compared to the topics in cluster 1, where not only Belarus but also the role of protests, the importance of human rights, strikes, and, interestingly, the efforts towards the Energy Union are discussed. However, energetics is not included in this cluster or among the most active countries regarding tweeting behavior. The representative tweet of the most important topic (Topic 4) revolved around precisely the role of energy in the political crisis of Belarus, primarily through the discussion about Russia's efforts to bypass the traditional transit countries. The Yamal–Europe gas pipeline has been connecting Russian natural gas fields in the Yamal Peninsula and Western Siberia with Poland and Germany through Belarus since 1997 (Gazprom, 2020). Belarus has achieved solid economic growth despite a lack of natural resources and the economic crisis that followed the dissolution of the Soviet Union, primarily through manufacturing and exports of industrial machinery, mineral products, chemical products, metal alloys, and fabrics. Belarus, on the other hand, is heavily reliant on imports to fulfill its energy requirements.

Almost all electricity generation in 2018 (97 percent, or 39 terawatt-hours [TWh]) was derived from natural gas, but this is expected to revolutionize with the commencement of two nuclear power plants (1200 megawatts [MW] each, scheduled to come online in October 2020 and July 2021) (IEA, 2020). However, Belarus is heavily reliant on imports of all types of fossil fuels, most of which are supplied by Russia. As we have seen on the cluster map, Austria is also part of this cluster because Austria receives their gas need through Belarus and the Yamal–Europe gas pipeline. Although the Nord Stream pipeline provides natural gas from Russia, it runs under the Baltic Sea, thus bypassing traditional transit countries such as Belarus and Ukraine. Overall, the most crucial topic of Cluster 2 was its vulnerability to the natural gas that comes through Belarus.

The next topic is related to the first one because it discusses those counties that made veto the EU sanctions against Belarus. The Belarus sanctions of the EU were in doubt because Cyprus demanded action against Turkey. The topic refers to the collision between two unrelated foreign policy crises. At the same time as the Belarusian crisis, there were rising tensions in the eastern Mediterranean over the Turkish drilling to access gas reserves in the Mediterranean. The crisis started escalating when the French president, Emmanuel Macron, led those inside the EU to oppose Turkey's increasingly military foreign policy and say Turkey could no longer be seen as a partner in the Mediterranean. Macron offered Greek Prime Minister Kyriakos Mitsotakis French military aid, including the sale of 18 Rafale jets. The issue was raised at a meeting of the Med7 group of southern Mediterranean leaders on the French island of Corsica and again at an EU Council meeting on March 23. Overall they discussed the demand for access to much of the eastern Mediterranean. Germany, the lead mediator between Turkey and Greece, was exploring the possibility of expanding Turkey's customs union with the EU to resolve a dispute exacerbated by massive hydrocarbon discoveries in the eastern Mediterranean over the past decade. Turkey has long sought a broader customs union with the EU (Rodríguez, 2020).

The next topic discusses the situation of press situation and protests in Belarus. In Minsk, tweets discussed an extraordinary demonstration of Flower Power took place. During a challenging period for Belarus, women halted police abuse against citizens by standing with flowers along significant thoroughfares, demonstrating boundless love. They greeted all cars with flowers while dressed in white, and cars responded with signals. This peaceful demonstration against autocrat Lukashenko was dubbed the "Flower Revolution." (Makhovsky, 2020) Another part of this topic discussed the possibility that Moscow had an intention to send unmarked soldiers (Little Green Man) to Belarus to support Lukashenko's reign. The penultimate topic is closely connected to the previous one as tweets discussed the solidarity with protesters, the number of protesters, and the police brutality against the demonstrators. The last topic focuses on the different strikes in Belarus, such as the demonstration of the significant Belarus tractor factory where workers were threatened with dismissal. On Friday, 14 August 2020, protesters included workers from the Azot plant in Grodno, Minsk Automobile Plant, or the Salihorsk mine plant.

Figure 25: Topic modeling results for tweets in Cluster 2 (Belarus)

Table 11: Topic modeling results of Cluster 2 (East) with representative text for each topic (RQ2b & RQ3b)

(The text is already pre-processed so it might not always make sense grammatically)

Topic #	Topic Contribution %	Keywords	Representative Text
0	0.9717	lukashenko, russia, eu, putin, ukraine, sanction, russian, cyprus, turkey, country	denmark chides cyprus for blocking belarus sanctions
1	0.9687	president, election, eu, lukashenko, belarusian, call, human_rights, lithuania, opposition, member	belarus as neighbors belarus are appealed to the belarusian authorities to fully recognize and comply with basic democratic standards we are calling for refraining from violence and to respect fundamental freedoms, human rights and citizen including minority rights
2	0.9663	minsk, protest, lukashenko, people, woman, detain, police, today, protester, belarusprotest	bloody night in minsk cichoanous in fact a few days ago i did not know who this courageous woman who stood against the communist dictator belarus
3	0.9757	twiby, lukaschenko, art, artist, stopdictator, prayforbelarus_stopdictator, lukasenka, exhibition, prayforbelarus, helpbelarus	march for peace minsk belarus belarusfreedom belarussolidarity standwithbelarus supportbelarus prayforbelarus peace
4	**0.9796**	pan_european, european_gas, coal, interdependence_humanityfirst,	and what about belarus they are not worth to speak about them and also about ns II this is not regarding sovereignty where germany bought energy but only about

		energy_economic, oil_community, create_great, energyunion_greennewdeal, futureofeurope_covid, start_let	geopolitics
5	0.9619	child, company, strike, worker, employee, nexta, director, plant, radio, tractor	many are now talking about general strike to belarus it is a hyper industrial country the old assembly factory of the ussr tractors fridges televisions are yjs products the workers have a huge weight to follow
6	0.9698	journalist, weapon, woman, revolution, remove, use, school, flowers_come, airport, belarusian_flower	belarus lukashenko putin littlegreenmen flowerrevolution belarus invasive those little green men putin invasive forces making lebensraum ninabahinskaya the belarusian flower revolution where have the flowers come from the women have used them as weapons
7	0.9698	people, help, freedom, support, know, thank, go, world, want, live	belarusian people raised the flag of his freedom belarus belongs to be free independent democratic european country standwithbelarus belarus belarusprotests belarussolidarity freebelarus belarus

8.5 Evaluation of the results in the light of action logics

Due to the fact that our topic modeling approach is based on time series clustering, which identifies the most similar locations (in our case, countries) in a space-time cube and separates them into distinct clusters whose members share similar time series characteristics. We can analyze the existence and correlations of different constructing logics among the topics. It makes visible how these groups may influence the co-creation of mutual understandings comprising a community.

8.5.1 #Allforjan protest

Connective action

The logic of connective action can be distinguished in tweets categorised as sensitivity, as the collected tweets are emerging through the network, where the large number of unique actors in these large-scale fluid networks illustrates a movement without formal organizing. No central leadership, however, the process reflecting the shift from institutional (governmental or non-governmental) organization to more individual organizing. Initially, the local and international colleagues of Ján Kuciak represent the formal organizational actors who used social technologies, such as the creation and dissemination of the #AllforJan hashtag or the online publication and translation of Kuciak's final article, to link weak public networks and develop around individualized action themes (for instance, corruption or on-site, online protests). Therefore, the movement is organized using technical affordances, employing the hashtag's organizational capacity to self-organize massive networks of people into a spontaneously rising chorus of voices. This occurs via interactions with the network, that may include isolated people who use hashtags such as #AllforJan or #zaslusneslovensko (for a decent Slovakia) as a reference point. The action connective manner is further supported by the existence of five different spatial clusters. It also illustrates the broad territorial expansions, since technologies that serve as mediators have made it possible to connect a wide flexible network with no geographical limitations.

Although the colleagues of Kuciak may be considered as active agents with their early supporting activity, the movement itself was self-organizing. It is represented by the absence of a central leader or organization, and it is based on the free engagement of actors who actively take part in co-creation. We can see this in the large number of topics, which exemplify how individuals used personalized action frames to define and translate the #Allforjan movement, and thereby co-create the movement by adding their own unique insight. In the initial phase of the analyzed timeframe, Slovakia and Malta provided a peak at the same time because they had a similar case with the assassination of Daphne Caruana

Galizia. The general and customized action frames derived from this material had the characteristic to flow across geographical and cultural limitations.

Another manifested aspect of the connective action is the application of technology as the organizational tool in combination with the enormous number of individualised narratives provided via natural involvement. Others embrace the act of sharing these narratives with the aim of self-validation via likes and retweets, that serves as a self-motivating component. This is clear considering the increasing number of tweets and the collaborative tweets that often urge people to share and read #Allforjan tweets, follow them and join the protest movement or simply endorsing others by like or retweeting becomes self-expression and self-validation when it contributes to the collective good and serves as a legitimization phase. It reflected the capabilities of Twitter in this process, as Twitter-interactions rely heavily on others reciprocating this act of information exchange by retweeting, likeing or posting anyone else's content, enhancing the prospective impact of the movement as a tool for facilitating social change.

Collective action

This forging of a collective identity seems to have its origins in the drive to pursue revolutionary progress cooperatively by adopting a shared purpose and working together. In the various topics, they widened the collective identity by arguing that corruption and politicians are the problems (Cluster 1: Topic 5 and 6). This may form ties among the distant individuals. Moreover, the personal narratives are equally crucial to collective action since individualized action frames are articulated into collective action frames via ongoing encounters and cooperation among actors, creating a shared cause collectively (e.g., for a decent Slovakia, Cluster 5: Topic 6). The cross-creation of a collective identity is thus influenced by these collective frameworks (#allforjan). The abundance of exchanges that reflect the local citizens' collective identification showed this, as they are the opponents of the governing elite ("Fico must go")

The construction of collective action frames, based on the collective identity framing, is also identified in the general theme of the topics. Most of them attempt to highlight a perceived general understanding; which refers to a "we and they" divided structure: “#Allforjan because people are more concerned about why citizens stay in suppressive governance: “In the light of recent events democracy remains unimaginable without press freedom stresses publisher Detlef Prinz Jan Kuciak Daphne Caruana Galizia" (Cluster 5, Topic 4), „The slaughter of colleagues continues now it's the turn of Ján Kuciak in Slovakia the coordination for the safety of journalists set up in Italy is a model in the European states" (Cluster 1, Topic 1). These cases show the use of collective action framing in which journalists are victims of political corruption.

Another important identity frame may be lying in the overlapping representations of female victims. In the Topics (Cluster 1, Topic 1 and 4 or Cluster 5, Topic 0) may reflect a component of gender-based violence wherein women (Martina Kušnírová, Daphne Caruana Galizia) are perceived as subjects of oppression and a society that does very little to aid them. The various representations (fiancee, girlfriend or relative) of Martina Kušnírová among the Topics (Cluster 1: Topic 1, 3, 4; Cluster 5: Topic 0, 6) emphasized an aspect of violence against the private sphere. Either the aggression towards women or violence against the private sphere is both a consequence some level of inequality. This exemplifies how users are linked by similar value systems and beliefs that problematize this specific social phenomenon and how they take part in co-creation of the narrative, and thereby the co-construction of a collective identity („Slovaks for a decent Slovakia").

Community

The subsection that follows outlines a few of the most significant contents comprising the community and how these shared views as a system are affected by the logics of action. During our topic modeling approach, we considered the hashtags as an integral part of narrative, therefore the hashtags as the organizing agent appears in the various Topics thus revealing the group's mental construction.

The analysis process and effect of action logics show the co-creation of a mutual understanding and the significance of revealing personal views. Using customized action frames that are driven by the act of sharing to demonstrate a self-organizing network identifies a common knowledge of the impact of individual stories. The significance of personal narratives is in their ability to bring new perspectives, since these narratives based on customized action frames provide something unique, although the core event itself may be well-known, for instance: e.g. „our story [colleagues of Jan] was never completed due to Jan's murder, here is Part I why" (Cluster 1: Topic 5).

They also characterized the individual narratives as being employed in the path to self-validation, as the individuals who share via customized action frames seeking communal legitimacy; „absolutely sick Slovakia is truly going back to its post communists era rip to Jan Kuciak" (Cluster 1: Topic 5). This shows how personal narratives are communicated with the audience and illustrates the community's commitment. This may be determined by analyzing the topics of the shared narratives that explain the participants' experiences or perspectives on the event. These personal narratives are intriguing in terms of mutual trust, since these contacts facilitate the unification practice and, by extension, cooperation between actors. This shows the potential for encouraging interpersonal connections, cooperation, which may cause the formation of stronger bonds: "On Friday, 9th March 2018 at 17:00 we will march again we demand a new and trustworthy government.", which potentially inspires trust and express a broader belief.

This shared goal establishes the foundation for the building of a unified "we" (we colleagues, we journalists, we Slovakians) that is the outcome of ongoing contact and cooperation. This is essential for the formation of communities because it may provide emotional support and a sense of group identity.

Collaboration

Collaboration is essential because it may bring individuals who don't get along closer together and enable them to work together to show solidarity and create trust. We recognize collaboration as a common understanding within the community, since the community comprises individuals who strive to make sense of the world by comparing their perceptions, or personal experiences, with those of others (Brass et al., 2004). This practically means that individuals based on their need for a sense of connection involve in the co-creation of communities, e.g. "Moraes take part in mission to Slovakia [...] how can the EU better protect journalists in Europe?" (Cluster 5: Topic 0) or translating Kuciak's last article into Italian. The mutual understanding and cooperation depends on the inclination for voluntary engagement by persons who take part in co-creation via the self-motivated act of sharing, exemplifying the effect of connective action logic. Every actor who shares material and uses the same hashtag (#allforJan; #zaslusneslovensko; #korupcia) contributes to the co-creation of the community. The engagement of a diverse range of persons shows unconstrained access, demonstrating how weak-tied interactions connect social peripheries can facilitate the co-creation of high-quality user-generated material, that brings value to the community and may recruit and keep additional people.

Collaboration is also impactful on the act of sharing, since the action lacks formal structure and is network-dependent, making each connection essential. The potential inherent in a collaborative endeavor which relates to the collective action of establishing a common cause, thus energized individual connections. This also suggests that, through forming a common cause and collectively pursuing social change, the community may be recognized as a collaborative effort with a larger group of individuals whose involvement is normalized based on a few social signals.

Collaboration arises in individuals engaging, who are joined in this ad hoc networks based on a similar goal and, through collective effort, pursue change in society and co-creates the mutual understanding; it is necessary to increase attention. The shared notion brings awareness represents the formation of a common cause, since the necessity to increase awareness stems from the collective identity that journalists or even citizens may be the victims of a corrupt political system. This collective identity is continually co-created and sustained by actors' interactions, participation, and collaboration based on the perception of a single "we."

Summary

We identified how narratives depict connective action through the self-motivated action of sharing in the quest for reinforcement under the use of tailored action frames. In particular, we found personal narratives in the topics act as a medium for co-creation and collaboration and can foster confidence in the community and interactions, culminating in the development of stronger links. Furthermore, cooperation was recognized in the collaborative forging of a common cause, which encouraged the build of a collective identity, an coordinated "we," which induced supplementary trust in the society and urging collaboration for social change since a combined effort results in a decision. It was founded on the collective identity framing and the transformation of individualized action frames into communal action frames motivated by co-operation.

Overall, the #AllforJan movement has components of both action logics. This is perceptible by how the correlations among topics reflect the self-organizing network. The application of personalized action frames exemplified by connective action, and the collective co-creation of a shared cause that results in creating a centralized "we" as illustrated by collective action. Conseqently, the #AllforJan movement functioned as a hybrid network.

8.5.2 #Belarusprotest

Connective action

The connective logic may be characterized by the specific topics rising across the network as Tweets, illustrating a movement without formal leaders and organization. In the case of Belarus, Tsikhanouskaya shortly after the election results fled the country. However, before it she was a solid leader of the opposition in place. The transformation of the movement in Belarus may stick to this event. Indeed, seeing the interaction amongst protesters (see timeline): the lack of rallies, the protest marches as climax, are certainly a remarkable manifestation of the Belarusian events' nature. The timely regular but individual marches are a noticeable characteristic of the protest wave's most significant activities. Its effect has been to highlight types of peaceful demonstrations in which all participants seem to be equal, such as the various marches (march of unity, march of women or students or seniors) or human chains. Walking in a mass march formation to a designated endpoint is definitely an effective form of a protest, however without a meeting at beginning, or rally, to hear speakers it may lose its power. As we have seen prior, the sentiment distribution of Cluster 2 aligns with this pattern, as the majority of Monday's most optimistic values appear after midnight, reflecting the protest marches but the intermittent time remained pessimistic, while the baseline of the mood constantly decreased from the middle of August 2020.

Initially, the opposition represented the formal organizational actors who used social technologies, such as the creation and dissemination of various hashtags (relating to protest informations) to link weak public networks and develop around individualized action themes. However, the local government partially shut down internet access of multiple sites for local citizens with the assistance of Belarus’s National Traffic Exchange Center, which manages the Belarus’ internet networks, which may have made difficult to exploit the benefits of Internet. Using different technical tools supported the flow of the movement, to exploiting the organizational capacity to self-organize networks of people. In the Cluster 2, Topic 5 reveals a channel of information flow which was provided by the Nexta, a Belarusian media site that is mainly dispersed Telegram and YouTube channels.

Using social media inevitably led to interactions inside the network, but the usage of different channels fragmented its impact. It is shown by the limited number of spatial clusters that may essentially be divided into two groups, East and West. Besides the spatial component, another limitation element is the lack of unifying hashtags outside the Cluster 2 (East). The word Belarus as hashtag in the Cluster 1 (West) bear not any additional meaning while in Cluster 2 (East) contains hashtags like belarusprotest, prayforbelarus, helpbelarus which may increase a consciousness of shared interests and aims and sympathies while creating a psychological sense of unity among the groups and classes, but it could not break out from East (Cluster 2). Overall, these characteristics emphasize the geographical limitations of the network.

The Belarusian protest movement was self-organized, since a number of entities may be regarded as active agents due to their supportive activities. It is characterized by the lack of a central leader or organization and is founded on the free involvement of individuals in co-creation. This is clear in the broad range of subjects, which illustrate how people employed personalized action frames to define and interpret the protest, co-creating the movement by contributing their own distinctive perspective. Cluster 2, Topic 5 reveals the relevance when the workers of the Minsk Tractor Factory (MTZ), a major Belarus plant, went on strike in protest against police violence following the country's disputed election. Another vital characteristic of the movement is the representation of women actors in the Cluster 2, Topic 2 (courageous woman, cichoanous) and 6 (belarusian_flower, woman) which could slightly seep into the Cluster 1 (Topic 2).

Collective action

This forging of a collective identity seems to have its origins in the drive to pursue revolutionary progress cooperatively by adopting a shared purpose and working together. The appearance of the word cichoanous (Cluster 2, Topic 2) reveals a unique aspect of the movement, as the word may reflect Chicana feminism, an American feminist movement. Chicana feminists questioned their prescribed status in their family and claimed further acknowledgment of their experiences. They identify as self-aware, self-reliant, and proud of their origins, traditions, and experiences, while placing the community (La Raza) in the first place of their life. With the rise of the Movement, their family structures underwent profound transformations. Specifically, women examined the pros and cons of the existing family dynamic and their role in the Chicano national fight (Garcia, 1989). This may create connection between distant individuals through the network.

The active role of the women is the major topic in the construction of collective action frames, based on the collective identity framing. Indeed, one of the most noticeable aspects of the protest movement in Belarus is the disproportionate participation of women. A woman, Sviatlana Tsikhanouskaya, has emerged as the improbable opponent to the long-serving president of Alexander Lukashenko of Belarus. Two of the most prominent opposition activists in the nation who have been kidnapped or forced to depart the country are women. One of them, Maria Kolesnikova, ripped up her papers to avoid being exiled against her will. And women led some of the initial protests against the election outcome.

The ability to connect with personal narratives to collective action frame is crucial since individualized actions may transform into collective action frames via ongoing encounters and cooperation among actors. In Belarus, however, gender norms are strongly maintained by social pressure. Women are expected to assume the role of homemaker and continue to stay inside the family sphere in order to protect and sustain the traditional family structure. Moreover, Belarus lacks both institutional and legislative frameworks for women's equality, despite the fact that there are no legal obstacles to women's involvement in administration. Based on the statistics from 2014, Belarus has just one female minister, and none of the seven chairs of regional committees are women. Although the "we" (women) and "they" (man) division in the political setup can create the collective action frames, its success may be limited, especially in similar administrative circumstances that Belarus has.

Community

Even before the exit polls, Belarus was the already a site of minor protests. The demonstrations did not turn into a broad phenomenon until the analytical phase, which coincided with Lukashenko's victory announcement. Although men and women took part in demonstrations, men substantially in number were more than women in the early days. However, the government attempted to stop the spread of the movement by force. The pattern of early protests shows the disproportionate and unjustified use of force against citizens by police authorities. Cluster 2, Topic 2 describes the "Bloody Night in Minsk," which may refer to the social standards of Belarus. The social hierarchy, gender stereotypes, customs, and expectations also make demonstrators vulnerable, since it is more comprehensible to beat males. The Cluster 2 supports this notion as it does not mention serious injuries at all. As violence against males increased, women went to the streets in support. The sensation of vulnerability may strengthen emotional support and collective identification; yet, the violence against a particular group (males) lowers the power of the movement (gendered vulnerability), which ultimately led to the movement's transformation into a series of peaceful marches.

Colaboration

Collaboration is essential because it may bring individuals who don't get along closer together and enable them to work together to show solidarity and create trust. We recognize collaboration as a common understanding within the community, since the community comprises individuals who strive to make sense of the world by comparing their perceptions, or personal experiences, with those of others. This practically means that individuals based on their need for a sense of connection involve in the co-creation of communities. The existence of collaboration can be identified in both Cluster and the heatmap, too. Cluster 1, Topic 5 shows the European Union restrictive measures against Belarus, while Cluster 2, Topic 1 and the heatmap along the whole period reveals Latvia and Estonia supporting role. The mutual understanding and cooperation depends on the inclination for voluntary engagement by persons who take part in co-creation via the self-motivated act of sharing, exemplifying the effect of connective action logic.

Summary

We distinguished how topics portray connective action through the self-motivated action of sharing in the pursuit of reinforcement using a partially customised action frames. Among the subjects, we did not identify a significant number of personal accounts. Despite the identification of co-creation and cooperation, the number of clusters and the topic's association revealed the restriction of movement in Belarus. They acknowledged the collaborative forging of a shared purpose and fostered the development of a communal identity within a well-defined geographical region. Despite the fact that the #Belarus movement has elements of both action logics, both have limitations that hindered the movement's effectiveness. This is evident by the low correlations across topics, which are reflective of the self-organizing network. Nonetheless, the Belarusian movement operated as a mixed network.

Table 12: Comparsion of the case-studies in the light of action logics

	#AllforJan	#Belarusprotest
Organization	A technologically-based self-organizing network that generates a decentralized movement.	A technologically-based self-organizing network that generates a decentralized movement.
Structure	Extensive flexible network with unrestricted network access and interactions. Connections are created with the movement, we have also observed although social relationships and deeper bonds at the interpersonal level.	Extensive flexible network with partially restricted network access and interactions. Connections are created with the movement, we have observed, although social norm shapes the bonds at the interpersonal level.
Motivation	The action of self-motivated sharing served as self-validation and encouraged spontaneous engagement. The act of self-motivated sharing served as self-validation and encouraged voluntary engagement. The social norm legitimized engagement and cooperation by motivating others to	Individuals step-by-step engaged in continual interactions with the social norm, which served as a sign linking dispersed individuals by communicating solidarity and unity.

	participate and contribute to the common good.	
Identity framing	The establishment of a common objective encouraged the formation of a single "we" rooted in a common believe in the narrative. This resulted in the formation of a shared identity.	The establishment of a common objective encouraged the formation of a single "we" rooted in a common believe in the narrative. However, the social norm enabled only a partially validation of the shared identity.
Conversational content	In the interactions and self-motivated acts of sharing and involvement, personalized action frames were apparent, indicating the self-organizing character.	In the interactions and self-motivated acts of sharing and involvement, personalized action frames were apparent, indicating the self-organizing character.
Limits	No obvious limit other than the creation of a common cause based on emotional reactions and the gravity of the issue, which may need an actor's dedication if they adapt.	Participation needed either offline action or participation online as the form of commitment. Authorities applied some constraining considerations.

8.5.3 A unique difference between Slovakia and Belarus

The tragic events of death were present in both of the case studies that we examined. On the other hand, the way they came across on their social media profiles was very different. According to the news report, five men were killed in Belarus as a result of police brutality during the time period that was under consideration. Their names are Alexander Taraikovsky, Alexander Vikhor, Artsyom Parukou, Konstantin Shishmakov, and Hienadz Shutau. The topic modeling approach, on the other hand, did not bring to light these additional events. If we dig a little deeper, we can also discover that these topics were not a base component of the communication that was tied to the demonstration. In contrast to Belarus, funerals have emerged as a visible indicator of growing levels of protest in Slovakia. Despite the fact that numerous historical movements (Biggs, 2005; Alexander, 2006; Alexander, 2011) demonstrate that the function that images play is by no means insignificant, a large number of earlier studies have neglected how visual representation are conveyed and how they effect individuals.

Visual representation can symbolize and compress events and ideologies. According to Goldberg, in modern society, these representations are signs (Goldberg, 1993) that may appear in the form of the representation of drastic deaths, as happened in several countries during the Arab Spring before the protests began. On 6 June 2010, young Khaled Saeed's life tragically ended in Alexandria, Egypt; shortly after that, two faded police officers dragged the 28-year-old out of the local internet cafe while he was beaten so severely that he succumbed to his injuries within hours. Pictures of his injuries were spreading unstoppably on various social media surfaces. On 17 December 2010, the life of a certain Mohamed Bouazizi, a Tunisian street vendor, ended similarly tragic, but in his case, the sole cause was the brutality that ended his own life. She set herself on fire in front of the local government building after an officer publicly dishonored and confiscated her shares. Shocking, outrageous images of the man on fire and his hospitalization spread across the internet. In May 2011, a similar incident happened with a 13-year-old boy in Dara, Syria. Hamza Ali Al-Khateeb was arrested during a protest in late April at Dara, from which only his lifeless body was returned to his family with severe injuries, including burns and gunshots. His post-mortem pictures spread a similar intensity in the online sphere as the previous examples.

The prior study of the social media groups formed during the various revolutions has shown that the pre-revolutionary websites acted as a gathering group, allowing users to expose and share the brutality of the police and other bodies while at the same time helping organize protests and later these groups focused on democratic transformation (Khamis & Vaughn, 2011). Concerning news examination, Papacharissi and de Fatima Oliveira found that the content of violence made up more than half of the news. Similarly, quantitative analysis of social media posts has shown that their number may vary between different phases of the protest, especially during a significant event. Overall, the number of shared content on social media correlates with on-site events (Freelon, 2011). However, this type of communication can only be abstract, a set of information that appears in writing or orally. The other branch of communication is the emotional one, which bypasses the built-in rationality of language and thus directly affects the viewer's moral senses. Therefore, there is reason to believe that attention-grabbing visual themes can serve as indicators for demonstrations.

Neurological research shows that the human brain perceives visual stimuli more effectively than verbal cues because they are closely linked to survival, allowing them to attract more attention. From the point of view of evolutionary psychology, the processing of messages with harmful content in particular triggers similar processes in the human brain as indirect threats, which is why it could be essential to investigate the early impact of images in social movements. We try to examine and understand the role of images in exploding demonstrations. How a photograph, which often depicts a moment of death or other sufferings, becomes symbolic (Gazzaniga, 2005). A photo of suffering or death is primarily personal. It can only transform itself into a universal symbol of injustice if it not only receives more publicity besides the private sphere, but this publicity can also interpret the additional information behind the picture.

Visual representation may have a high emotional charge. It is one thing to read about atrocity and another to experience its consequences visually, notably when it comes to the human body. Images of physical suffering play a central role in this regard and may assistance produce what Jasper and Poulsen call “moral shocks” (Jasper & Poulsen, 1995). The sight of bodily suffering is powerful because the body symbolizes the inevitable fate for everyone. Regarding the representation of bodily fragility, Linfield believes that the cruelty displayed by an image shatters the person’s perception of what it means to be a human being (Linfield, 2011). Therefore, viewing violent images elicits emotional reactions, and emotions can affect human behavior.

The importance and impact of attention-grabbing content become clear when we look at the motivators of political participation. While Susan Sontag was skeptical of the transformative power of images depicting suffering, social psychological research suggests that emotion, along with rationality, predicts an individual's willingness to take collective action. According to Sontag, images that emotionally shock the viewer rather than mobilize them can be more demobilizing (Sontag, 1978). This view is contradicted by social psychological findings, which argue that anger encourages individuals to organize against systemic inequalities, indirectly suggesting that images of physical brutality can influence viewers' emotional responses and lead to collective action.

Two possible factors that increase the efficiency of the social movement or protest are the size and level of identification with the movement. As we have seen, intermediary networks allow users to use personalized frameworks that allow participants to share social and political issues that interest them (Bennett & Segerberg, 2014). These digitally mediated, personalized frames have proven faster, more mobilized, and flexible. In this system, social movement activists try to support or act against emotions appearing on the web to keep the mobilization going (Goodwin & Pfaff, 2001). Given the motivating potential of emotions and their proliferation in online communities (Stage, 2013), there is no doubt that images and their emotional content may have played a role in erupting the movement.

A key predictor of willingness to take collective action is how effective an individual feels in engaging in a collective effort to achieve social change. According to explanations of efficiency, the size of a social movement lies in the fact that it shapes the perception of individuals. According to Klandermans, willingness to participate is critical in determining the number of participants in collective action (Klandermans, 1984). According to Van Zomeren, the possible social support, i.e., the expected number of participants, determines the movement's effectiveness. Feedback on this number is given on social media when members of a group see the number of members (van Zomeren et al., 2004). Identification with the group also motivates people. Thus, both the number of people in the group and the identification with the movement increase the group members' collective perception of effectiveness and promote participation in the social movement.

In a repressive society, the main obstacle to change and political expression is the individual's fear of being disadvantaged by his or her actions. Social media effectively help overcome this fear by removing the sense of isolation in cyberspace. At the same time, he emphasizes that the individual is also an integral, creative, and valuable part of the community, as individuals with the same values and aspirations but with territorial isolation can unite in any corner of the online space. Social media are thus able to demonstrate the collective character of a movement, to strengthen dissidents' sense of belonging to the community. Thus, a growing group in cyberspace can engage and motivate more and more individuals and empower them to take action through the nature of the crowd.

However, a representation cannot be a symbol on its own, but only with the participation of activists (Schlegel, 1995). The image can be loaded with content that fits into a particular society's political and cultural system to influence it, and it becomes part of the community's memory, shaping that society. The advantage of photographs is that they are linked to reality based on similarities, which allows them to depict perfect reality over other branches of art (Hans et al., 1978). Although all photographs can be interpreted within the existing framework, the most crucial component in the development is the context and politics through which activists create a symbol. So the picture's reality depends on the viewer's knowledge, so it can only gain meaning in the proper context. If this does not happen, the picture is nothing more than a picture.

However, if the image is widely interpretable, it may become a universal symbol for a significant part of society. At the core of this process are the activists creating universally interpretable knowledge and interpretation behind the image, thus forcing the respective society to self-reflection. The activists, therefore, create an interpretation in the existing political context that is crucial for the transformation of the image into a symbol. Since photographs alone cannot establish a moral position, through their abstract compression characteristic, they can reinforce an existing interpretation or develop a new one. Activists created an interpretive framework that supports the moral principles of regaining or transforming power. The interpretation may symbolize a broader spectrum of injustice when connected to an existing interpretive framework. The dynamic interaction between image and society can therefore only take place with the participation of activists.

The contributions of activists were similar to the cases of Saeed and Bouazizi, and Al-Khateeb. After the family was notified of Saeed's death, they had to identify him officially. At this point, Saeed's brother took a photo using his cell phone of his brother's injuries, showing traces of blows to his face: bruises and a broken and deformed jawbone. He then uploaded the photos of Saeed's injuries to his own social media page and an earlier passport photo depicting the once vibrant young man. This comparison further highlighted the brutality of the police. Bahaa Eltawel, Saeed's family acquaintance, who worked as a journalist at Youm7, saw the juxtaposed images on social media. On 6 June, Eltawel published an article about Saeed's death. He contrasted the circumstances of his death with police reports and later widened the contrast by obtaining documents testifying Saeed's impunity and exemplary military service, thus undermining the credibility of the official explanation. The article containing the photos was noticed by Wael Ghonim, an Egyptian employee at Google who created a Facebook page called "We are all Khaled Said" on 8 June 2010, where he shared the article and the images. The site has become a significant gathering group of young people of a similar age, dissatisfied with the system, where they shared their thoughts on the case and the domestic political situation in Egypt in general (Ghonim, 2012). On 14 January 2011, the group announced a demonstration called "Revolution against Torture, Corruption, Unemployment, and Injustice" on 25 January, which took the group to the streets. Eventually, it became the first demonstration of the movements that later caused the system's collapse.

The photograph of Saeed’s death and its information were already widely known when Bouazizi set himself on fire at the Governor’s Office on 16 December 2010, after he was not allowed to complain. The incident took place near a busy road next, in front of the eyes of many, many of whom were trying to help the blazing Bouazizi. 90% of his body was burned. He was transported first to the local hospital, then to Sfaxi Hospital, and later to the Ben Arousi Burn Center, where he finally died on 4 January 2011. More than 5,000 people attended his funeral, and his tombstone was decorated with the following inscription: “Martyr Mohamed Bouazizi” (Chesshyre, 2013). However, the protests did not begin after his death but hours after his suicide, as Bouazizi’s action took place in public, so pictures of the incident were in the news minutes later. The protests gradually intensified, eventually forcing President Zine el-Abidine Ben Ali to flee the country with his family to Jeddah, Saudi Arabia.

Months later, on 25 May 2011, 13-year-old Hamza Ali Al-Khateeb died in Daraa custody due to police brutality. The family sent photos and videos to various newspapers showing the young man's injuries. Like Saeed, a Facebook page was created for him called We Are All Hamza Alkhateeb, which had more than 100,000 users by the end of May and featured face-to-death pictures. In this case, his symbolism also took place after his death became known: on 31 May, Hillary Clinton called his death a turning point, symbolic of many other situations in Syria (Sly, 2011).

The previous pages showed three examples of modern movements that have emerged with the help of social media. The cases of Saeed, Bouazizi, and Al-Khateeb represent deep-rooted social and political issues that caused a moral shock to society with their deaths and portrayals, but only after activists molded them into symbols. The images were transformed into symbols in two steps in all three cases. First, before and after death photos became simultaneously available on social media sites by juxtaposing them, highlighting the morally shocking nature of the after death photos. Then when additional information was added to the pictures, various news reports, or family involvement. Of course, this is not a new historical phenomenon, but the emergence and spread of the Internet, social networks, and global communication networks have facilitated the emergence and spread of these symbols in the region, which have thereby reinforced each other.

Although the analyzed case studies of this dissertation contained tragic deaths, their online representation was not the same. Although the deaths during the protests in Belarus appeared in the news, there was not able to transform into a symbol. In Slovakia, in contrast, the event had a strong emotional context. The young couple planned to marry in May 2018; however, on the forenoon of 26 February, the family called the police to the couple's home because they had not answered their phone calls for more than four days. The unfortunate event caused shock across the nation and mistrust in the government, sparking mass protests and a political crisis. The day after the first report was published, gatherings were held around the country in tribute. The media focused on Ján Kuciak's role in the tragic events because he worked as a journalist at the Aktuality.sk news site, where he focused mainly on investigating the tax fraud of several business people with connections to Slovakian politicians. However, the fiancée's funeral was held a day before the fiancé's. Martina Kušnírová was buried in her wedding dress on Friday, 2 March, in a ceremony attended by hundreds of mourners in Gregorovce. The heatmap that analyzed the Slovakian events revealed a high peak at the same das. Shortly after the event, a journalist issued the #allforjan hashtag on his Twitter page, then in the evening, up to 25,000 people protested in Bratislava against the attacks. On 9 March, protests were held in 48 cities in Slovakia and 17 other towns around the world. In Bratislava alone, about 60,000 people held a protest march. Two weeks later, the series of events culminated in the resignation of the Prime Minister and his cabinet.

As we have shown in earlier chapters, protest movements can also emerge within a diverse combination of numerous countries: a democratic state, a strategic regime that considers the world as a source of tools and resources and a familiar dictatorship in which we feel comfortable in our surroundings. Such obstacles frequently inspire curiosity, causing individuals to view their environment with new perspectives. Lastly, there is always the aspect of justification: moments of protest are also times of vigorous civic discussion about the common good — the principles from which the demonstration stems, and not merely the tools necessary for fighting for those objectives and values. Focusing only on the interests and assets through a densed topic modeling can reveal much less about a social setting than understanding the interrelationships between the many components. Coordination during a protest involves more than merely discussing who does what and when in order to accomplish a particular outcome. It also involves harmonizing the claims of the various engagement regimes that a given scenario requires. The trail is well-known. It begins when people in authority invade our most intimate environment, and in the most emblematic cases, the human body: Bouazizi (Tunisia), Al-Khateeb (Egypt), and the Kuciak-couple. The victims become potent symbols of protest that represent the innumerable identical encroachments faced by a large number of people; other symbols represent items with equal intrinsic value. The subsequent protest movement develops modes of participation that are discussing the common good at issue, searching for resources to right wrongs, questioning old societal assumptions, and exploring different ways to relate to others. This phenomenon has been observed in Slovakia, whereas Belarus emphasizes an alternative approach.

DISCUSSION

Sociological questionnaires, whether conducted on-site or online, have been an essential tool for comprehending group dynamics and the general social phenomena. Might the growing use of various social media sites create the risk of displacement of questionnaires? The large quantity of data that can be obtained from social media sites such as Twitter makes such sources enticing. A limited amount of research has attempted to utilize these sources pragmatically in order to comprehend or explain individual or group behavior from a social science perspective, while the questionnaires still a significant instrument for developing this knowledge. For decades, surveys have been an important data gathering tool for both policymakers and researchers. Recent developments in social media have offered a new growing data source and a new point of view from which to evaluate individual and collective behavior. While the volume of social media data has led some to promulgate the end of conventional survey devices, several in the social sciences have objected to the scientific usefulness of social media on the grounds of generalization and accuracy.

Probably the reality lies somewhere between these two ends of the spectrum. The social media and conventional survey work might complement one another to generate a deeper, more immediate portrayal of society. However, using social media, scientists may investigate real behavior around high-profile events, albeit online, as opposed to survey respondents' self-reported assumptions or intentions. Social media data collection is faster and less expensive than conventional surveys. Traditional surveys, on the other hand, may give deeper understanding into demographics and the objectives of individuals, as well as higher relevance to the current study issue. Given these advantages, disadvantages, and similarities, the value of social media in the social sciences remains to be discovered. This thesis has presented a complicated technique for analyzing the benefits and drawbacks of social media in comparison to conventional survey instruments.

The current work provided an in-depth analysis of the tweets posted in Europe related to the assassination of Kuciak and the after-election protests in Belarus. Twitter has many advantages as a data source in analyzing collective actions, especially in terms of the high temporal resolution thanks to the people’s immediate response considering given events and news, which we intended to illustrate with our multilayered analysis. Although there are no similar methods that can analyze social phenomena such as protests and reactions to such events at that large scale and with sufficient temporal resolution, the present methodology is also not free from limitations. It is well-known that Twitter is not representative of the whole population in terms of demographics; however, we can also hypothesize that in the case of a protest or other follow-up events, elderly or very young generations are also not very likely to be highly active or involved.

We analyzed two distant cases, so we cannot and should not draw general conclusions from the patterns we identified both in terms of spatiotemporal characteristics of the topics and the sentiments. Still, we can state that based on the news report and other official sources, our analysis is able to reflect what was going on in the countries we considered. Additionally, the workflow itself, containing the whole pre-processing and the follow-up analysis steps to cluster the countries and investigate topics and sentiment patterns in the data, can be considered transferable and used to analyze similar cases. This exploratory analysis can provide a strong foundation for more specific analyses, for example, country-specific investigation or focusing on a specific topic or day. Although we applied the current methodology based on Twitter data *after* a particular event had already taken place, it is worth noting that both the subject and the region of the analysis are interchangeable with other themes, as well as with other social media data resources that have the required attributes regarding the location. Moreover, by using unsupervised topic modeling methodologies and an automatable approach, there is also the potential to adapt our workflow for early classification (for example, in the emerging phase of a protest) to predict the overall gravity of the analyzed collective actions.

CONCLUSIONS

This research proposes a comprehensive technique to overcome the limits of current methodologies and to deal with the complexity of protest studies. The suggested method integrates multilingual corpus translation, location and sentiment extraction, and machine-learning topic modeling techniques to expose the underlying interests and motivations behind collective behavior. Consequently, the given method has a major advantage over previous research that focused heavily on hashtag activism (while disregarding geographical dimensions) or, alternatively, used solely location-specific hashtags. In contrast, by using machine learning methods and methodologies that may be nearly totally automated, we can get a considerably broader variety of input data than in present studies, in which researchers manually review all submissions.

This study investigated the patterns and similarities of Twitter data's geographical, temporal, and emotional indicators by building a fresh data-driven integrated methodology to investigate two unique East European protest movements. First is the 2018 European impact of the murders of Slovakian journalist Ján Kuciak and his fiancée Martina Kusnirova (26 February and 15 March 2018). Second is the impact of the 2020 Belarusian presidential election (9 August–23 September), which has not been studied by current research.

The multi-spectral interpretation of the complex and dynamic nature of the events prompted an effective way of analysis to facilitate in a more thorough comprehension of the protest dynamics. Consequently, our work outperforms the state of the art (SOTA) in two separate ways:

1. This thesis showed that georeferenced social media data may be utilized to analyze political events, even on a smaller geographical and societal scale and in languages other than English.
2. The current study proposes a novel algorithmic approach that integrates time-series clustering with semantic topic modeling and sentiment analytics on geo-referenced social media data in the analytics of social phenomenon.

By presenting an original perspective and considering the presented limitations of existing analyses, this research put effort into profoundly understanding the relevant research of the past decade. The specific results of the research were described in the Results chapter of the dissertation over fifty pages. However, we would like to emphasize again that both the literature review and the methodological setup also provide a unique contribution of the research. Below, I briefly summarize the most important results, item by item:

- **Increased geoparshing accuracy**
  According to the Twitter Developer Platform, approximately 1-2% of Tweets are geo-tagged, while ~30-40% of Tweets contain some profile location information. Our Belarusian dataset had 655,423 non-empty rows in the user_location field and 10,353 places in the user_bio field. The described geoparsher method was able to find 346,162 places in the user_location field and 10,353 places in the user_bio field. The idea behind analyzing this field was that users tend to use it to indicate their occupation and their employer. Overall, the presented mixed-model geolocating approach was able to identify 356,515 geolocated places, which is 54,3% of the Belarusian dataset. The application of the same method on the #Allforjan dataset returned coordinates for 8069 tweets, which is 61.2% of the original dataset. Kapanova and Stoykova (2019) conducted a comparable study and were able to analyze approximately 29 geolocated Tweets per day, while our methodology provided an average of 100 for each day for Slovakia. This is a 15–20% improvement on average compared to what we knew about the Twitter Developer Platform and a 244% increase compared to their similar study.

- **Increased semantic understanding of the data**
  Prior sociological studies on social media has primarily focused on hashtag activity, overlooking the written language, which is a crucial component of human communication (e.g, LeFebvre & Armstrong, 2018, Sinpeng, 2021). This dissertation focuses on the text, including multi-lingual corpus translation and sentiment extraction, utilizing machine-learning topic modeling techniques to expose the underlying interests and motivations of collective action. Consequently,

our technique has a distinct advantage over previous studies, which often ignored these factors.

- **Better understanding the spatial factor in the dynamics of contemporary protest movements**
  In order to comprehend spatial aspects in group dynamics, sociological study on social media has depended primarily on hashtag activity, which provide limited understanding. To comprehend the escalation of demonstrations based on social media activity, the methodology of this thesis did spatiotemporal analysis using geolocated tweets (see above). The data was aggregated at the nation level due to the possibility that a country's linguistic and political factors may impact the movement that was evaluated. The subsequent clustering algorithm identifies nations with comparable tweeting patterns, which may also suggest when demonstrations occurred or the existence of other influential characteristics, such as media and politics. By selecting countries with a similar pattern, our method was able to highlight the contributing characteristics and how they evolved through time. Moreover, it gave a more concise and illuminating visualization and explanation of the result than statistical values for the countries of Europe, one by one.

- **The general spatial characteristics with advanced visualization revealed the supportive countries of the #Allforjan movement**
  The spatial outcomes of the thesis's multilayered spatial analysis demonstrated a direct association with European countries. Although the absolute activity statistics indicated that Slovakia, Germany, Italy, and France were the most active nations, our method eliminated the bias caused by the impact of population sizes. Normalization with each country's population and improved visualization uncovered two crucial factors: active countries and outliers. Belgium, the Czech Republic, Germany, France, Hungary, Poland, Slovakia, and Switzerland were the eight countries with consistent tweeting activity during the course of the 18 days evaluated in our analysis. Czechia, Hungary, and Poland are three neighboring countries with a shared past. Three outliers were shown by the visualization: Malta,

Italy, and Slovakia. A few months earlier to the assassination of Kuciak, a journalist from Malta, Daphne Caruana Galizia, was also murdered. This piqued Malta's curiosity. Due to the relationship between Italian 'Ndrangheta mafia and Slovakian officials, the Italian's position was widely debated throughout the period. Switzerland was the publisher of the periodical for which Kuciak was employed. The normalized heatmap depiction indicated the three most significant theme groups that contributed to the establishment of the protest network.

- **The general temporal characteristics revealed the crucial days of the #Allforjan movement**

  One of most notable spike occurred on 28 February, when Kuciak's incomplete study about the connections between the Italian mafia and Slovakian politicians was revealed, while there are other lesser peaks beginning on the main protest day (9 March) and the resignations of Slovakia's interior minister and prime minister (12 and 14 March).

- **Clustering of these countries based on similarity revealed the dynamics the main characteristics of the #Allforjan movement**

  Our algorithm made five groups: Cluster 1 (blue) had high tweeting activity (1 March) in the first few days and a second, smaller peak when PM Fico resigned (14 March). Italy has the most tweets in this group, mainly because of its mafia's involvement in Slovakia's murders and corruption. Cluster 2 (red) had the least Kuciak-related tweets. Cluster 3 (green) has a modest activity level as Cluster 2, with two smaller peaks on 28 February and 14 March. Cluster 4 only includes three countries, and there's no peak when the murder and motivation were uncovered. 12 and 15 March are peaks for these countries due to the PM's resignation and other indirect effects of the murder or journalist's work. If we look at the subject modeling results of various countries, we can see that tweets about the murder and subsequent events may have a political narrative embedded in their political systems. The trend is obvious, but the absolute quantity of tweets is low (similar to Cluster 3), suggesting Twitter may not be the most popular social media tool in these countries. Those that use it may be less representative of the broader population than in other nations, where more tweets were harvested, distorting the results. Clusters 2–4 lack enough tweets for a deep study of sentiment patterns and subject modeling. Cluster 5 (purple) has over 5500 tweets and includes Slovakia. The group peaked when the journalist was found dead and his unfinished work was published. The number of tweets dropped to its lowest point on 8 March, before rising again on 9, 12, and 14 March.

- **Sentiment distribution per countries revealed the feedback of the #Allforjan movement and make insight of identity framing**
  Considering the government's ties to the mafia, most people supported the claims and were satisfied with the political consequences, such as the PM's resignation. Using the initial obtained sentiment compound score (before applying sentiment classes), we observed that Slovakia and Germany are among the nations with at least one tweet per day in the analysis period, a statistically rising trend. (Germany had 95% confidence, Slovakia 99%)

- **The topic modeling of the most active clusters (1 and 5) revelealed the main drivers of the #Allforjan movement and the weight of other political factors**
  Cluster 1 includes the PM's departure, the Italian mafia, and the EU's role. Most tweets denounce the murder, followed by concerns about press freedom and security. Kuciak's article, released after his death on February 28 in English and Slovakian, later appeared in other languages (e.g., in French). The article showed links between Slovakia's aristocracy and organized crime. The remaining discussions debate Kuciak's paper (Ndrangheta in Slovakia). Cluster 5 topic modeling findings show 7 subjects. These subjects are clearer Cluster 2 themes discussed not only the PM but also the government's role, protests, the mafia, Kuciak's fiancé, and Viktor Orban, the Hungarian PM, although Hungary was neither in this cluster nor among the most active countries in terms of tweeting behavior. Topic 2's representative tweet discussed the political turmoil in Slovakia, including resignations and new elections. The next topic covers press freedom in Europe, and its urgency is emphasized by the fact that several tweets claimed a clear connection between the deaths of Kuciak and Daphne Caruana Galizia, a

Maltese investigative journalist who was assassinated previously, on 16th October 2017. This theme's overrepresentation (Topics 4 and 0) demonstrates that Caruana Galizia's case bolstered the Kuciak movement online. The tweets that linked Galizia's death to Kuciak's fiancée's case might be viewed as a clear denunciation of violence against women that may fuel this campaign. Martina Kusnirova is mentioned by name and as a fiancée, suggesting that users of this cluster not only lamented her death but also differentiated between an innocent and a work-related fatality. Topic modeling also identified tweets on the Hungarian PM and Hungarian-born American billionaire George Soros, a common issue in conspiracy theories and fake news. This may have two causes. First, six months before Kuciak's killing, the Hungarian government initiated a statewide billboard campaign depicting George Soros and declaring, "Don't let George Soros have the last laugh," making him a scapegoat for the 2016 refugee crisis. It may have helped Slovakian PM Fico's 5 March 2018 political declaration. In this statement, the PM asked about George Soros and Andrej Kiska, who had proclaimed new elections a day earlier. PM Fico may have discredited the president with this statement. Second, PM Orban may have seen George Soros' "fingerprint" in the Slovakian crisis on 10 March. This portrayal of Soros may reflect similar political inclinations among different countries, as the clustering algorithm put Hungary, Belarus, and Turkey in the same group.

- **The general spatial characteristics with advanced visualization revealed the supportive countries of the #Belarusprotest movement**

  Multilayered spatial analysis in the research showed a link between the protest and other European countries. Our strategy eliminated the bias created by population size, which revealed that the UK, Germany, Belarus, and the Russian Federation were the most active nations. Normalizing population and improving visualization revealed active countries and outliers. 33 countries have continuous tweeting during our 44-day analysis. Among these, Western neighbors Latvia, Lithuania and Estonia were the most important ones. These nations first sanctioned Belarus. The Lithuanian parliament imposed economic sanctions on Alexander Lukashenko and 30 Belarusian officials on August 18. These countries also incorporated the judgments of the European Union. The analytics revealed an outlier, namely Austria where the narrative was stick to a possible energetic crisis.

- **The general temporal characteristics revealed the crucial days of the # Belarusprotest movement**
  One of most notable spike occurred when Konstantin Shishmakov, 29, director of Vaukavysk's Bagration Military History Museum, disappeared on 15 and 16 August 2020. As a deputy of the election committee, he refused to accept and sign the forged documents. He called his wife at 5 p.m. and said he was coming home, but he never arrived. On 18 August, Shishmakov's body was found in a river in Grodno. From 15 August, smaller peaks appear. Most likely, these are the biggest protest days because of the "solidarity chain" marches, which protested the crackdown in Belarus following the election and the police violence that followed, leading to multiple deaths and arrests. Josep Borrell, the EU's High Representative, said on August 14 in Brussels that the EU would prosecute Belarusian officials who rigged the election and used violence. On 16 and 23 August, Belarus's largest protest days, a "March of Unity" was held in Minsk with 200,000 participants and all major regional centers. 6,000 people protested in Homel, 4,000 in Hrodna, and 3,000 in Brest, Vitebsk, and Mogilev, which may explain why they were discussed across Europe. Sviatlana Tsikhanouskaya, the leading opposition candidate against Alexander Lukashenko, escaped to Lithuania after the election results were published.

- **Clustering of thes countries based on similarity revealed the dynamics the main characteristics of the #Belarusprotest movement**
  The most imporant result of the clustering is the numbers of the clusters which is two, revealing the spatial limitation of the movement. The first cluster (blue) covers European nations with strong tweeting activity in the first week. After the first peak (18 August), it displays a diminishing pattern with smaller peaks on 20, 24, 28, and 31 August, as well as 7, 14, and 21 September. In addition, tension rises between 7 and 9 September. Mondays have single peaks, indicating Sunday protests. Cluster 1 includes wise political decisions and occurrences. The first turning point came

when state-controlled firms joined the demonstrators on 18 August. The opposition created the Coordination Council that day to aid a power transfer. The Belarusian chief prosecutor started a criminal probe into the Council two days later (20 August). Multiple global events affected web trends on August 24. First, publications helped spread the news that 50,000 Lithuanians joined human chain protests yesterday. Tsikhanouskaya met the U.S. deputy secretary as Belarusian police detained two Coordination Council members. EU foreign ministers resolved to penalize 20 Belarusian officials on August 28. On August 31, Coordination Council member Lilia Vlasova was detained. Between 7 and 9 September, Cluster 1 users focused on protest leaders who were kidnapped and carried to the Ukraine border but unable to leave. The 1st and 2nd clusters' trends diverge significantly until 24 August. Cluster 2 (red) Eastern countries had the least active users. In these countries, tweeting activity is modest and falling, with three smaller peaks on 24, 31 August, 7, 14, 21 September. Cluster 2 peaked a day early than Cluster 1 because of demonstrations. On the fifth day of the demonstration, 13 August, participants formed huge lines. If we look at the topic modeling results of Cluster 2, we can see a high possibility that tweets will mention protests, their participation, insults, and the rally's impact.

- **Sentiment distribution per countries revealed the feedback of the #Belarusprotest movement, revealing a certain distance from the events on the part of Western countries**

  Cluster 1 is an fluctuation of positive and negative values. This cluster's users viewed various September events positively, especially on 11, 15, 19, and 21. These were mostly political developments. The EU denounced the violence and demanded Belarus free all jailed demonstrators. The most unfavorable qualities were the largest protests and their aftermaths in Belarus, such as the long solidarity queues (13 August) and when three senior opposition organizers went missing a day after 10,000 Belarusians marched through Minsk. The tweets show a notable focus on protest days, Sundays and Mondays. Protest days dominate tweets. The majority of Monday's positive numbers occur after midnight, indicating protest marches. Users tend to be negative in the middle, especially a day before protests, which shows how they feel about the government. Cluster 1 analytics show that most respondents supported the charges, given Lukashenko's election fraud and Europe's response. Overall, European users were satisfied with political outcomes like the Coordination Council and EU sanctions but nothing more. Cluster 2 (East) has two negative extremes on 13 August and 7 September. Cluster 1 sentiment changed more dramatically than Cluster 2. Cluster 2 users were greatly affected by the protests, whereas Belarusian factory workers' countrywide strike was regarded good. Cluster 2 users concentrated on negative incidents in Belarus, such as a self-immolation attempt near a police station in Smaliavichy or when 390 women were held by Belarusian soldiers on 19 September. Cluster 2 countries and users focused on how Belarusians live and suffer, while Cluster 1 focused on the shifting political scenario (for example, by putting sanctions on Belarus or seeing the Coordination Council as a legitimate power).

- **The topic modeling of the most active clusters revealed the main drivers of the #Belarusprotest movement together with political and limiting factors**
  The issues of Cluster 1 (West) include protests in Minsk, Alexander Lukashenko's relationship with Vladimir Putin, and EU sanctions against Belarus. Overall, leaders' tweets condemning events in Belarus were the most important. Emanuel Macron said Belarus leader Lukashenko "must go," and Joe Biden chastised Trump for not condemning the despot. Lukashenko's amicable relationship with the Russian president and its effect on events and Belarus' neighbors follow. Cluster 1 focused on European politics and Belarus. Cluster 2 (East) subjects have higher percentage contributions than West topics (Cluster 1). These subjects include protests, human rights, strikes, and the Energy Union. Topic 4's representative tweet focused on the significance of energy in Belarus' political turmoil, namely Russia's efforts to bypass traditional transit countries. The Yamal–Europe gas pipeline connects the Yamal Peninsula and Western Siberia to Poland and Germany through Belarus. Austria is included since it gets its gas from Belarus and the Yamal–Europe gas pipeline. Cluster 2's biggest issue was its reliance on Belarusian natural gas. Women's role in the protest is another important topic. Tweets discussed a Flower Power demonstration in Minsk. Belarusian women stopped police abuse by

standing with flowers along major thoroughfares, showing unlimited love. They greeted drivers with white flowers and signals. "Flower Revolution" was the peaceful protest against tyrant Lukashenko. Overall, there is no interaction and multiole connection between Cluster 1 and 2.

- **The unique difference between Slovakia and Belarus was the handling of deaths which revealed a limitation factor for Belarus**
  Both case studies included people who tragically died, yet their online reputations were absolutely different. During the protests in Belarus, five men were killed by police brutality. In contrast, these supplementary occurrences were not revealed by the topic modeling strategy. Further investigation reveals that these points were not crucial to the discussion surrounding the presentation itself. A obvious indicator of rising protest levels in Slovakia was funerals, contrary to Belarus.

- **The comparison of the events in Slovakia and Belarus in the light of connective and collective logics revealed those factors that led success only in Slovakia**
  Both movements used technology to create a decentralized network. Their structures differ greatly. Slovakia had a large, flexible network with unlimited access. Movement fostered connections and deeper interpersonal bonds. In Belarus, the government partially controlled network access and exchanges, so the extensive, flexible network was not fully integrated. Slovakia's social norm legitimized engagement and cooperation by encouraging others to contribute to the common good. Belarus' societal norms may be limiting. Personal narratives must be connected to collective action frames because individual actions can become collective action frames through continual encounters and cooperation. Social pressure in Belarus maintains gender norms. Belarus lacks institutional and statutory foundations for women's equality, despite no legal barriers to women in administration. The "we" (women) versus "they" (men) division in politics might establish collective action frames, but its success may be limited, especially in Belarus.

9 #Allforjan datasets

Table 13: Detailed data of #AllforJan dataset)

	#tweets	#users	avg_twt_cnt	5_twt_usr	5_twt_cnt	5_twt	max_twt	max_daily_avg	rest_usr	cnt_over_avg	pcnt_over_avg	stdev_twt_cnt
Slovakia	1876	468	4,008547	86,97%	640	34,12%	8,21%	8,55	57,68%	74	15,81%	10,26611
Germany	1115	457	2,4398249	90,81%	621	55,70%	6,01%	3,72	38,30%	93	20,35%	4,5571
Italy	825	403	2,0471464	94,29%	553	67,03%	6,30%	2,88	26,67%	67	16,63%	3,58868
France	635	356	1,7837078	96,35%	492	77,48%	5,04%	1,77	17,48%	93	26,12%	2,41155
Belgium	621	224	2,7723214	91,07%	307	49,44%	18,52%	6,38	32,05%	43	19,20%	8,07039
Czechia	606	252	2,4047619	91,67%	325	53,63%	5,45%	1,83	40,92%	46	18,25%	4,08768
United Kingdom	415	251	1,6533864	98,01%	340	81,93%	6,02%	1,33	12,05%	57	22,71%	2,32674
Poland	350	127	2,7559055	89,76%	173	49,43%	14,29%	2,77	36,29%	27	21,26%	5,73113
Spain	263	122	2,1557377	96,72%	150	57,03%	35,36%	5,16	7,60%	13	10,66%	8,32734
Switzerland	239	89	2,6853932	92,13%	120	50,21%	16,74%	2,22	33,05%	16	17,98%	5,47023
Austria	161	80	2,0125	95,00%	122	75,78%	11,80%	1,05	12,42%	14	17,50%	2,43153
Netherlands	140	95	1,4736842	96,84%	120	85,71%	5,71%	0,44	8,57%	23	24,21%	1,15021
Malta	92	35	2,6285714	88,57%	55	59,78%	11,96%	0,61	28,26%	10	28,57%	2,58662
Sweden	49	36	1,3611111	97,22%	43	87,76%	12,24%	0,33	0	8	22,22%	0,91750
Serbia	42	28	1,5	96,43%	36	85,71%	14,29%	0,33	0	5	17,86%	1,26773

Denmark	35	16	2,1875	93,75%	19	54,29%	45,71%	0,88	0	1	6,25%	3,59198
Greece	27	13	2,0769230	92,31%	21	77,78%	22,22%	0,33	0	4	30,77%	1,73034
Russian Federation	27	14	1,9285714	92,86%	20	74,07%	25,93%	0,38	0	4	28,57%	1,70981
Albania	22	6	3,6666666	83,33%	5	22,73%	77,27%	0,94	0	1	16,67%	5,96284
Bosnia and Herzegovina	18	7	2,5714285	85,71%	8	44,44%	55,56%	0,55	0	1	14,29%	3,11022
Moldova, Republic of	15	6	2,5	100,00%	15	100,00%						0,76376
Luxembourg	10	7	1,4285714	100,00%	10	100,00%						0,49487
Latvia	4	4	1	100,00%	4	100,00%						0
Iceland	3	2	1,5	100,00%	3	100,00%						0,5
Lithuania	1	1	1	100,00%	1	100,00%						0

10 #Belarus datasets

Table 14: Detailed data of #Belarusprotest dataset)

	#tweets	**#users**	**avg_twt_cnt**	**5_twt_usr**	**5_twt_cnt**	**5_twt**	**max_twt**	**max_daily_avg**	**rest_usr**	**cnt_over_avg**	**pcnt_over_avg**	**stdev_twt_cnt**
Austria	5018	1160	4,32586	87,07%	1731	34,50%	4,48%	5,11	61,02%	191	16,47%	13,3793
Belarus	26908	2412	11,15588	83,29%	3299	12,26%	12,61%	77,11	75,13%	227	9,41%	86,7357
Denmark	2045	410	4,9878	89,27%	585	28,61%	44,65%	20,75	26,75%	61	14,88%	26,7904
Lithuania	2267	480	4,72291	86,88%	723	31,89%	6,44%	3,32	61,67%	75	15,63%	25,2706
Poland	8837	1742	5,07290	86,51%	2376	26,89%	9,09%	18,25	64,03%	235	13,49%	24,4326
Russian Federation	17496	3315	5,27782	87,81%	4367	24,96%	6,77%	26,91	68,27%	404	12,19%	28,4613
Ukraine	7296	1290	5,65581	85,43%	1744	23,90%	5,10%	8,45	71,00%	188	14,57%	20,6410
Belgium	9640	1516	6,35883	83,64%	2145	22,25%	13,94%	30,54	63,81%	215	14,18%	43,1628

Czechia	2891	723	3,99861	88,38%	1006	34,80%	15,91%	10,45	49,29%	127	17,57%	18,9163
Finland	3816	1114	3,42549	89,23%	1594	41,77%	5,87%	5,09	52,36%	200	17,95%	10,7301
France	14092	3377	4,17293	88,10%	4623	32,81%	3,87%	12,38	63,33%	486	14,39%	17,3359
Germany	39639	9457	4,19149	87,96%	13590	34,28%	2,34%	21,07	63,38%	1378	14,57%	17,2799
Ireland	2713	762	3,56036	91,99%	1051	38,74%	35,72%	22,02	25,54%	108	14,17%	35,2056
Italy	7333	2332	3,14451	91,30%	3276	44,67%	7,77%	12,95	47,55%	355	15,22%	14,0682
Latvia	7068	475	14,88	87,16%	713	10,09%	64,45%	103,52	25,47%	20	4,21%	211,0790
Netherlands	11664	2478	4,70702	89,75%	2224	19,07%	13,97%	37,02	58,98%	322	12,99%	40,0981
Portugal	2176	528	4,1212	91,29%	763	35,06%	15,26%	7,54	49,68%	45	8,52%	20,6624
Spain	8500	2968	2,86388	91,58%	4085	48,06%	4,16%	8,14	47,78%	611	20,59%	10,3230
Sweden	7294	2141	3,40681	89,63%	2947	40,40%	7,88%	13,07	51,71%	352	16,44%	14,9180
Switzerland	4195	993	4,22457	87,31%	1384	32,99%	8,96%	8,55	58,05%	148	14,90%	16,6976
United Kingdom	41019	12789	3,20736	91,53%	17677	43,09%	2,59%	24,18	54,31%	1801	14,08%	17,4822

Variables

#tweets	number of tweets
#users	number of users
avg_twt_cnt	average tweet count (#tweet/#users)
5_twt_usr	% of users who tweeted no more than 5 tweets
5_twt_cnt	count of tweets from users who tweeted less than 5
5_twt	% of tweets from users who tweeted no more than 5 tweets
max_twt	% of tweets posted by the user with the most tweets
max_daily_a	# of tweets/day for the most active user

vg	
rest_usr	tweet count for users who tweeted more than 5 times (except user with the most tweet)
over_avg	count of users who tweeted more than average tweet count
pcnt_over_avg	% of users tweeting higher than average
stdev_twt_cnt	standard deviation of tweet count/user

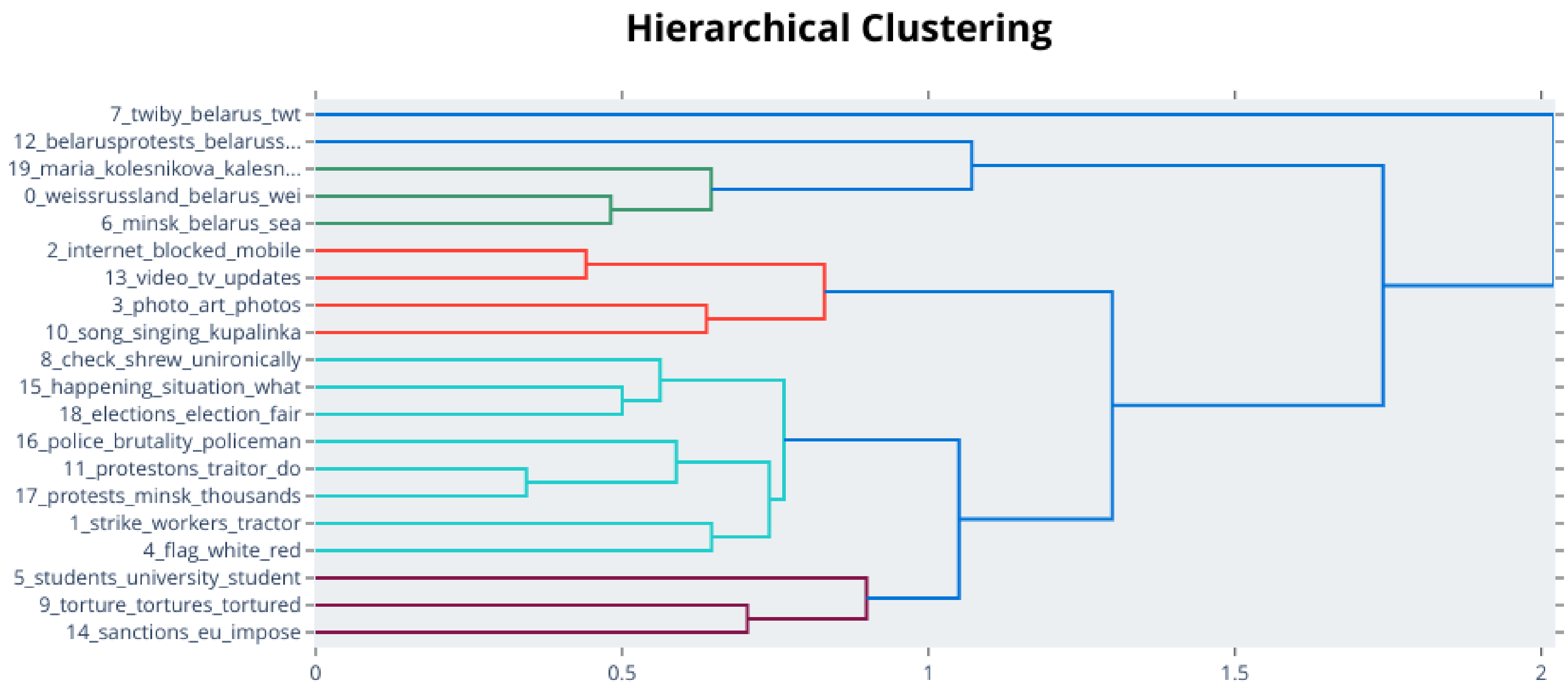


Figure 26: BERTopic hierarchical clustering of the topics (#Belarusprotest)

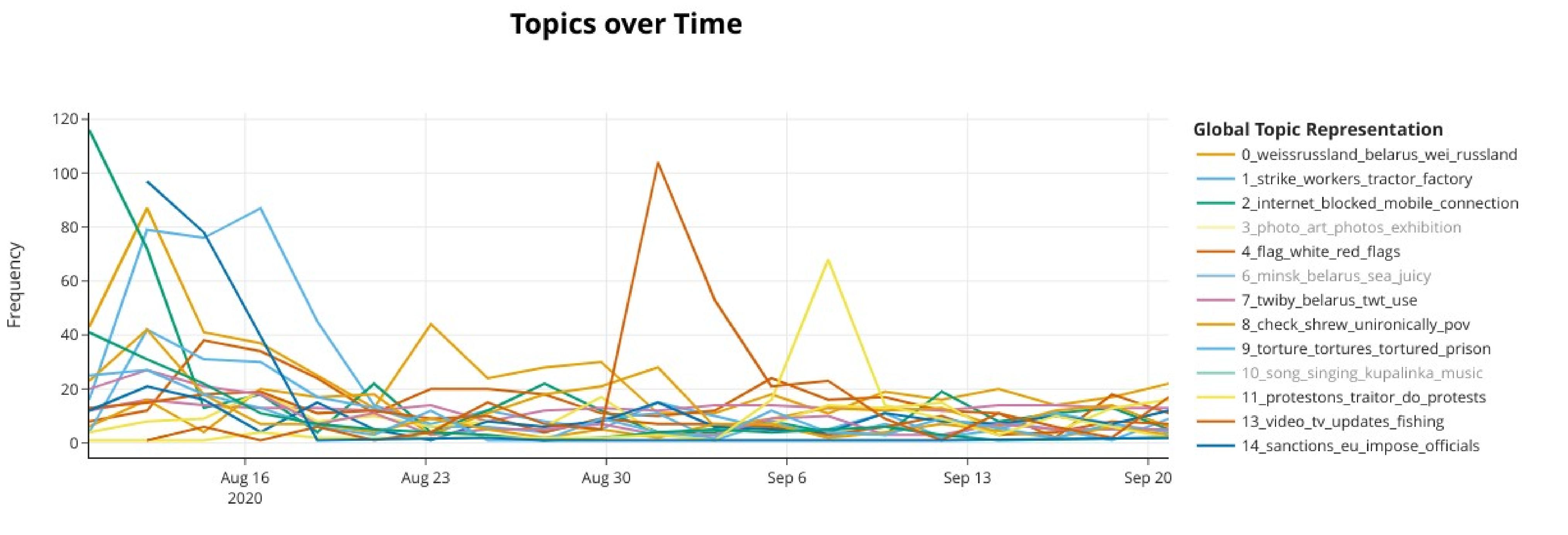


Figure 27: BERTopic topic tendencies over time (#Belarusprotest)

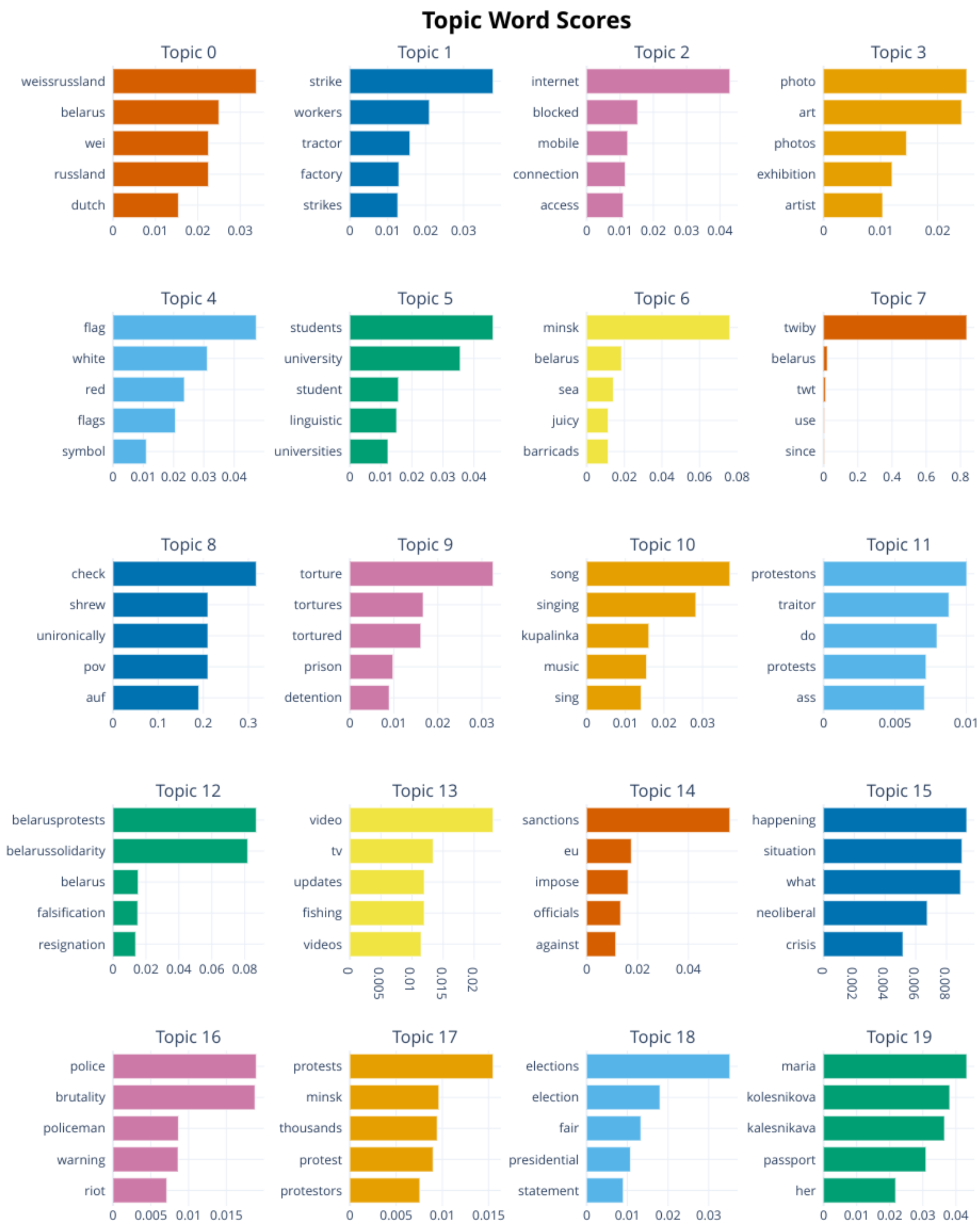


Figure 28: BERTopic topic word frequencies (#Belarusprotest)

BIBLIOGRAPHY

Abuzayed, A., & Al-Khalifa, H. (2021). BERT for Arabic Topic Modeling: An Experimental Study on BERTopic Technique. *Procedia Computer Science*, *189*, 191–194. https://doi.org/10.1016/j.procs.2021.05.096

Adams, P. C. (1996). Protest and the scale politics of telecommunications. *Political Geography*, *15*(5), 419–441. https://doi.org/10.1016/0962-6298(95)00088-7

Aday, S., Farrell, H., Freelon, D., Lynch, M., Sides, J., & Dewar, M. (2013). Watching From Afar: Media Consumption Patterns Around the Arab Spring. *American Behavioral Scientist*, *57*(7), 899–919. https://doi.org/10.1177/0002764213479373

Aiken, M. (2019). An Updated Evaluation of Google Translate Accuracy. *Studies in Linguistics and Literature*, *3*(3), 253–260. https://doi.org/10.22158/sll.v3n3p253

Alexander, J. C. (2006). *The Civil Sphere*. Oxford University.

Alexander, J. C. (2011). *Performance and Power*. Polity.

Alexe, A. (2018, January 18). New anti-corruption protest announced for January 20 in Bucharest. *Business Review*. https://business-review.eu/news/new-anti-corruption-protest-announced-for-january-20-in-bucharest-156203

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, *31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Alsaedi, N., Burnap, P., & Rana, O. (2017). Can We Predict a Riot? Disruptive Event Detection Using Twitter. *ACM Transactions on Internet Technology*, *17*(2), 1–26. https://doi.org/10.1145/2996183

AlSayyad, N., & Guvenc, M. (2015). Virtual Uprisings: On the Interaction of New Social Media, Traditional Media Coverage and Urban Space during the 'Arab Spring.' *Urban Studies*, *52*(11), 2018–2034. https://doi.org/10.1177/0042098013505881

Aslam, S. (2018, January 1). *Twitter by the Numbers (2018): Stats, Demographics & Fun Facts*. https://www.omnicoreagency.com/twitter-statistics/

Aust, C. F., & Zillmann, D. (1996). Effects of Victim Exemplification in Television News on Viewer Perception of Social Issues. *Journalism & Mass Communication Quarterly*, *73*(4), 787–803. https://doi.org/10.1177/107769909607300403

Bansal, P., Bansal, R., & Varma, V. (2015). Towards Deep Semantic Analysis of Hashtags. In A. Hanbury, G. Kazai, A. Rauber, & N. Fuhr (Eds.), *Advances in Information Retrieval* (Vol. 9022, pp. 453–464). Springer International Publishing. https://doi.org/10.1007/978-3-319-16354-3_50

Barron, L., Gunia, A., & Leung, H. (2019, October 1). *Hong Kong Rocked by Worst Unrest in Decades*. Https://Time.Com/. https://time.com/5690681/hong-kong-crisis-unrest-protests/

Barrons, G. (2012). 'Suleiman: Mubarak decided to step down #egypt #jan25 OH MY GOD': examining the use of social media in the 2011 Egyptian revolution.

Contemporary Arab Affairs, *5*(1), 54–67. https://doi.org/10.1080/17550912.2012.645669

Barry, A. M. (2005). Perception theory. In K. Kenney, K. Smith, S. Moriarty, & G. Barbatsis (Eds.), *Handbook of visual communication: Theory, methods, and media* (pp. 45–62). LEA.

Bastos, M. T., Mercea, D., & Charpentier, A. (2015). Tents, Tweets, and Events: The Interplay Between Ongoing Protests and Social Media: Tents, Tweets, and Events. *Journal of Communication*, *65*(2), 320–350. https://doi.org/10.1111/jcom.12145

Bastos, M. T., Recuero, R. D. C., & Zago, G. D. S. (2014). Taking tweets to the streets: A spatial analysis of the Vinegar Protests in Brazil. *First Monday*, *19*(3), Article 3. https://doi.org/10.5210/fm.v19i3.5227

Baumgarten, B., Daphi, P., & Ullrich, P. (2014). *Conceptualizing Culture in Social Movement Research*. Palgrave Macmillan.

BBC. (2009, July 6). Bulgaria opposition wins election. *BBC*. http://news.bbc.co.uk/2/hi/8134851.stm

BBC. (2011, October 24). England rioters “poor and young.” *BBC News*. //www.bbc.co.uk/news/uk-15426720

BBC. (2012, January 19). Bulgaria bans shale gas drilling with “fracking” method. *BBC*. https://www.bbc.com/news/world-europe-16626580

BBC. (2013, May 13). Bulgaria election fails to end political stalemate. *BBC*. https://www.bbc.com/news/world-europe-22498433

Bee, C., & Chrona, S. (2017). Youth activists and *occupygezi*: Patterns of social change in public policy and in civic and political activism in Turkey. *Turkish Studies*, *18*(1), 157–181. https://doi.org/10.1080/14683849.2016.1271722

Benedicto, J., & Kriesi, H. (Eds.). (1992). *Las transformaciones de lo político*. Alianza Ed.

Benford, R. D., & Snow, D. A. (2000). Framing Processes and Social Movements: An Overview and Assessment. *Annual Review of Sociology*, *26*, 611–639.

Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University.

Bennett, L. (2005). Social Movements beyond Borders: Organization, Communication, and Political Capacity in Two Eras of Transnational Activism. *Transnational Protest and Global Activism*, 203–226.

Bennett, W. (2003). Communicating Global Activism. *Information, Communication & Society*, *6*(2), 143–168. https://doi.org/10.1080/1369118032000093860a

Bennett, W. L., & Segerberg, A. (2012). The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics. *Information, Communication & Society*, *15*(5), 739–768. https://doi.org/10.1080/1369118X.2012.670661

Bennett, W. L., & Segerberg, A. (2014). *The logic of connective action: Digital media and the personalization of contentious politics*. Cambridge University.

Bennett, W. L., Segerberg, A., & Walker, S. (2014). Organization in the crowd: Peer production in large-scale networked protests. *Information, Communication & Society*, *17*(2), 232–260. https://doi.org/10.1080/1369118X.2013.870379

Berrett, D. (2011, October 16). *Intellectual Roots of Wall St. Protest Lie in Academe*. The Chronicle of Higher Education. https://www.chronicle.com/article/intellectual-roots-of-wall-st-protest-lie-in-academe/

Bhavnani, R., & Donnay, K. (2012). Here's Looking at You: The Arab Spring and Violence in Gaza, Israel and the West Bank: Violence in Gaza, Israel and the West Bank. *Swiss Political Science Review*, *18*(1), 124–131. https://doi.org/10.1111/j.1662-6370.2012.02056.x

Biggs, M. (2005). Dying without killing: Self-immolations, 1963-2002. In *Making sense of suicide missions* (pp. 173–208). Oxford University Press.

Bimber, B. A., & Davis, R. (2003). *Campaigning online: The Internet in U.S. elections*. Oxford University Press.

Bimber, B., Flanagin, A. J., & Stohl, C. (2005). Reconceptualizing Collective Action in the Contemporary Media Environment. *Communication Theory*, *15*(4), 365–388. https://doi.org/10.1111/j.1468-2885.2005.tb00340.x

Bimber, B., Stohl, C., & Flanagin, A. (2009). Technological change and the shifting nature of political organization. In A. Chadwick & P. Howard (Eds.), *Routledge Handbook of Internet Politics* (pp. 72–85). Routledge.

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python* (1st ed). O'Reilly.

Bliss, C. A., Kloumann, I. M., Harris, K. D., Danforth, C. M., & Dodds, P. S. (2012). Twitter reciprocal reply networks exhibit assortativity with respect to happiness. *ArXiv:1112.1010 [Physics]*. http://arxiv.org/abs/1112.1010

Blumler, J. G. (1970). The Political Effects of Television. In J. D. Halloran (Ed.), *The effects of television*. Panther.

Braun, E. (2020, September 27). *Macron: Lukashenko has to go*. Politico. https://www.politico.eu/article/emmanuel-macron-alexander-lukashenko-must-step-down-inauguration-illegitimate/

Bruns, A., & Burgess, J. (2012). Researching News Discussion on Twitter. *Journalism Studies*, *13*(5–6), 801–814. https://doi.org/10.1080/1461670X.2012.664428

Bruns, A., Highfield, T., & Burgess, J. (2013). The Arab Spring and Social Media Audiences: English and Arabic Twitter Users and Their Networks. *American Behavioral Scientist*, *57*(7), 871–898. https://doi.org/10.1177/0002764213479374

Buchanan, L., Bui, Q., & Patel, J. K. (2020, July 3). Black Lives Matter May Be the Largest Movement in U.S. History. *The New York Times*. https://www.nytimes.com/interactive/2020/07/03/us/george-floyd-protests-crowd-size.html

Buettner, R., & Buettner, K. (2016). A Systematic Literature Review of Twitter Research from a Socio-Political Revolution Perspective. *2016 49th Hawaii International*

Conference on System Sciences (HICSS), 2206–2215. https://doi.org/10.1109/HICSS.2016.277

Bush, D. (2020). *"Fighting Like a Lion for Serbia": An Analysis of Government-Linked Influence Operations in Serbia*. Stanford Internet Observatory.

Butler, J. (2009). *Frames of War: When Is Life Grievable?* Verso.

Cage, S., & Tsolova, T. (2013a, February 20). Bulgarian government resigns amid growing protests. *Reuters*. https://www.reuters.com/article/us-bulgaria-government/bulgarian-government-resigns-amid-growing-protests-idUSBRE91J09J20130220

Cage, S., & Tsolova, T. (2013b, February 28). Bulgaria president calls May election after protests. *Reuters*. https://www.reuters.com/article/us-bulgaria-government/bulgaria-president-calls-may-election-after-protests-idUSBRE91R0BO20130228

Campbell, A. (2021). What is Black Lives Matter and what are the aims? *BBC News*. https://www.bbc.com/news/explainers-53337780

Campbell, J. C., Hindle, A., & Stroulia, E. (2015a). Latent Dirichlet Allocation. In *The Art and Science of Analyzing Software Data* (pp. 139–159). Elsevier. https://doi.org/10.1016/B978-0-12-411519-4.00006-9

Campbell, J. C., Hindle, A., & Stroulia, E. (2015b). Chapter 6 - Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data. In C. Bird, T. Menzies, & T. Zimmermann (Eds.), *The Art and Science of Analyzing Software Data* (pp. 139–159). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-411519-4.00006-9

Cantor, N. F. (2021). *Age of Protest: Dissent and rebellion in the twentieth century.* ROUTLEDGE.

Castañeda, E. (2012). The *Indignados* of Spain: A Precedent to Occupy Wall Street. *Social Movement Studies*, *11*(3–4), 309–319. https://doi.org/10.1080/14742837.2012.708830

Castells, M. (2015). *Networks of Outrage and Hope: Social Movements in the Internet Age*. John Wiley & Sons.

Çelebi, A., & Özgür, A. (2018). Segmenting hashtags and analyzing their grammatical structure. *Journal of the Association for Information Science and Technology*, *69*(5), 675–686. https://doi.org/10.1002/asi.23989

Chadwick, A. (2007). Digital Network Repertoires and Organizational Hybridity. *Political Communication*, *24*(3), 283–301. https://doi.org/10.1080/10584600701471666

Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford University.

Charbel, J. (2011, January 22). Tunisia disproves Arab world's greatest myth. *Egypt Independent*. https://www.egyptindependent.com/tunisia-disproves-arab-worlds-greatest-myth/

Chen, S., Chen, S., Wang, Z., Liang, J., Yuan, X., Cao, N., & Wu, Y. (2016). *D-Map: Visual analysis of ego-centric information diffusion patterns in social media*. 41–50. https://doi.org/10.1109/VAST.2016.7883510

Chesshyre, T. (2013). *A Tourist in the Arab Spring*. Bradt Travel Guides.

Choi, S., & Park, H. W. (2015). Networking Interest and Networked Structure: A Quantitative Analysis of Twitter Data. *Social Science Computer Review*, *33*(2), 145–162. https://doi.org/10.1177/0894439314527054

Choudhary, A., Hendrix, W., Lee, K., Palsetia, D., & Liao, W.-K. (2012). Social Media Evolution of the Egyptian Revolution. *Communications of The ACM - CACM*, *55*, 74–80. https://doi.org/10.1145/2160718.2160736

Connors, J. P., Lei, S., & Kelly, M. (2012). Citizen Science in the Age of Neogeography: Utilizing Volunteered Geographic Information for Environmental Monitoring. *Annals of the Association of American Geographers*, *102*(6), 1267–1289. https://doi.org/10.1080/00045608.2011.627058

Conover, M. D., Ferrara, E., Menczer, F., & Flammini, A. (2013). The Digital Evolution of Occupy Wall Street. *PLoS ONE*, *8*(5), 1–5. https://doi.org/10.1371/journal.pone.0064679

Coppock, A., Guess, A., & Ternovski, J. (2016). When Treatments are Tweets: A Network Mobilization Experiment over Twitter. *Political Behavior*, *38*(1), 105–128. https://doi.org/10.1007/s11109-015-9308-6

Corriere Della Sera. (2011, October 15). Indignati d'Italia in piazza a Roma «Uniti per il cambiamento globale»—Corriere Roma. *Corriere Della Sera*. https://roma.corriere.it/roma/notizie/cronaca/11_ottobre_14/indignados-roma-manifestazione-corteo-sabato-1901827771558.shtml?fr=correlati

Cresci, S., Di Pietro, R., Petrocchi, M., Spognardi, A., & Tesconi, M. (2015). Fame for sale: Efficient detection of fake Twitter followers. *Decision Support Systems*, *80*, 56–71. https://doi.org/10.1016/j.dss.2015.09.003

Croeser, S., & Highfield, T. (2014). Occupy Oakland and #oo: Uses of Twitter within the Occupy movement. *First Monday*, *19*(3), 1–13.

Dalakoglou, D. (2013). *The Movement and the "Movement" of Syntagma Square*. Society for Cultural Anthropology. https://culanth.org/fieldsights/the-movement-and-the-movement-of-syntagma-square

Dallison, P. (2020, September 14). *'A friend is in trouble': Lukashenko gets $1.5B loan from Putin*. Politico. https://www.politico.eu/article/alexander-lukashenko-1-5b-loan-vladimir-putin/

Davies, J. C. (1962). Toward a Theory of Revolution. *American Sociological Review*, *27*(1), 5–19. https://doi.org/10.2307/2089714

Della Porta, D., & Diani, M. (2006). *Social movements: An introduction*. Blackwell.

DeLuca, K. M. (2005). *Image Politics: The New Rhetoric of Environmental Activism*. Routledge.

Diani, M. (1992). The Concept of Social Movement. *The Sociological Review*, *40*(1), 1–25. https://doi.org/10.1111/j.1467-954X.1992.tb02943.x

Diani, M. (2011). Networks and Internet into Perspective1: Networks and Internet into Perspective. *Swiss Political Science Review*, *17*(4), 469–474. https://doi.org/10.1111/j.1662-6370.2011.02040.x

Dimov, I., & Fidanova, S. (Eds.). (2021). Timeline Event Analysis of Social Network Communications Activity: The Case of Ján Kuciak. In *Advances in High Performance Computing Results of the International Conference on "High Performance Computing" Borovets, Bulgaria, 2019* (pp. 118–131). Springer International Publishing : Imprint: Springer. http://link.springer.com/10.1007/978-3-030-55347-0

D'Monte, L. (2013, 0 19). *Swine flu's tweet tweet causes online flutter*. Https://Www.Business-Standard.Com/. https://www.business-standard.com/article/technology/swine-flu-s-tweet-tweet-causes-online-flutter-109042900097_1.html

Doerr, B., Fouz, M., & Friedrich, T. (2012). Why rumors spread so quickly in social networks. *Communications of the ACM*, *55*(6), 70–75. https://doi.org/10.1145/2184319.2184338

Doerr, N., & Teune, S. (2012). The imagery of power facing the power of imagery: Towards a visual analysis of social movements. In K. Fahlenbrach, M. Klimke, J. Scharloth, & L. Wong (Eds.), *The Establishment Responds: Power, Politics, and Protest since 1945* (pp. 43–56). Palgrave Macmillan.

Donk, W. B. H. J. van de, Loader, B. D., Nixon, P. G., & Rucht, D. (Eds.). (2004). *Cyberprotest: New media, citizens, and social movements*. Routledge.

Douai, A. (2014). "The Police and the Populace": Canadian Media's Visual Framing of the 2010 G20 Toronto Summit. *Canadian Journal of Communication*, *39*(2). https://doi.org/10.22230/cjc.2014v39n2a2710

Drüeke, R., & Zobl, E. (2016). Online feminist protest against sexism: The German-language hashtag #aufschrei. *Feminist Media Studies*, *16*(1), 35–54. https://doi.org/10.1080/14680777.2015.1093071

Earl, J. (2013). *Digitally enabled social change: Activism in the internet age*. Mit Press.

Earl, J., & Kimport, K. (2011). *Digitally Enabled Social Change: Activism in the Internet Age*. MIT Press.

Earl, J., McKee Hurwitz, H., Mejia Mesinas, A., Tolan, M., & Arlotti, A. (2013). This Protest Will Be Tweeted: Twitter and Protest Policing During the Pittsburgh G20. *Information, Communication & Society*, *16*(4), 459–478. https://doi.org/10.1080/1369118X.2013.777756

Edrington, C. L., & Lee, N. (2018). Tweeting a Social Movement: Black Lives Matter and its use of Twitter to Share Information, Build Community, and Promote Action. *The Journal of Public Interest Communications*, *2*(2), 289. https://doi.org/10.32473/jpic.v2.i2.p289

Efron, M. (2011). Information Search and Retrieval in Microblogs. *JASIST*, *62*, 996–1008. https://doi.org/10.1002/asi.21512

European Comission. (2021, March 3). *Belarus—Trade—European Commission*. https://ec.europa.eu/trade/policy/countries-and-regions/countries/belarus/

Fahlenbrach, K., Klimke, M., Scharloth, J., & Wong, L. (Eds.). (2012). *The Establishment Responds: Power, Politics, and Protest since 1945*. Palgrave Macmillan.

Felmlee, D. H., Blanford, J. I., Matthews, S. A., & MacEachren, A. M. (2020). The geography of sentiment towards the Women's March of 2017. *PLOS ONE*, *15*(6), e0233994. https://doi.org/10.1371/journal.pone.0233994

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96–104. https://doi.org/10.1145/2818717

Foot, K. A., & Schneider, S. M. (2006). *Web Campaigning*. MIT Press.

Freelon, D. (2011). The MENA protests on Twitter: Some empirical data. *Dfreelon.Org*. http://dfreelon.org/2011/05/19/the-mena-pro- tests-on-twitter-some-empirical-data/

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*(5), 349–358. https://doi.org/10.1037/0003-066X.43.5.349

Furet, F. (1981). *Interpreting the French Revolution*. Cambridge University Press ; Editions de la Maison des sciences de l'homme.

G. Almatar, M., Alazmi, H. S., Li, L., & Fox, E. A. (2020). Applying GIS and Text Mining Methods to Twitter Data to Explore the Spatiotemporal Patterns of Topics of Interest in Kuwait. *ISPRS International Journal of Geo-Information*, *9*(12), 702. https://doi.org/10.3390/ijgi9120702

Gallagher, R. J., Reagan, A. J., Danforth, C., & Dodds, P. (2018). Divergent discourse between protests and counter-protests: #BlackLivesMatter and #AllLivesMatter. *PloS One*. https://doi.org/10.1371/journal.pone.0195644

Gazprom. (2020, March 21). *Yamal—Europe*. https://www.gazprom.com/projects/yamal-europe/

Gazzaniga, M. S. (2005). *The mind's past*. Univ. of California Press.

Gerbaudo, P. (2015). Protest avatars as memetic signifiers: Political profile pictures and the construction of collective identity on social media in the 2011 protest wave. *Information, Communication & Society*, *18*(8), 916–929. https://doi.org/10.1080/1369118X.2015.1043316

Gerbaudo, P. (2017). Social media teams as digital vanguards: The question of leadership in the management of key Facebook and Twitter accounts of Occupy Wall Street, Indignados and UK Uncut. *Information, Communication & Society*, *20*(2), 185–202. https://doi.org/10.1080/1369118X.2016.1161817

Ghonim, W. (2012). *Revolution 2.0: The Power of the People is Greater Than the People in Power : a Memoir*. Houghton Mifflin Harcourt.

Gleason, B. (2013). #Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter. *American Behavioral Scientist*, *57*(7), 966–982. https://doi.org/10.1177/0002764213479372

Goldberg, V. (1993). *The power of photography: How photographs changed our lives*. Abbeville.

Goldstone, J. A. (2001). Toward a Fourth Generation of Revolutionary Theory. *Annual Review of Political Science*, *4*(1), 139–187. https://doi.org/10.1146/annurev.polisci.4.1.139

González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. (2013). Broadcasters and Hidden Influentials in Online Protest Diffusion. *American Behavioral Scientist*, *57*(7), 943–965. https://doi.org/10.1177/0002764213479371

Goodwin, J., & Pfaff, S. (2001). Emotion Work in High-Risk Social Movements: Managing Fear in the U.S. and East German Civil Rights Movements. In *Passionate Politics* (pp. 282–302). University of Chicago Press.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, *101*(Supplement 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101

Grootendorst, M., & Reimers, N. (2021). *BERTopic* (v0.9.4) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.4381785

Gurr, T. R. (2015). *Why Men Rebel*. Routledge.

Haffner, M. (2019). A place-based analysis of #BlackLivesMatter and counter-protest content on Twitter. *GeoJournal*, *84*(5), 1257–1280. https://doi.org/10.1007/s10708-018-9919-7

Halfmann, D., & Young, M. (2010). War Pictures: The Grotesque as Mobilizing Tactic. *Mobilization*, *15*, 1–24.

Halterman, A. (2017). Mordecai: Full Text Geoparsing and Event Geocoding. *The Journal of Open Source Software*, *2*(9), 91. https://doi.org/10.21105/joss.00091

Hamanaka, S. (2020). The role of digital media in the 2011 Egyptian revolution. *Democratization*, *27*(5), 777–796. https://doi.org/10.1080/13510347.2020.1737676

Hans, J. S., Barthes, R., & Heath, S. (1978). Image-Music-Text. *The Journal of Aesthetics and Art Criticism*, *37*(2), 235. https://doi.org/10.2307/429854

Harlow, S., & Johnson, T. J. (2011). The Arab Spring| Overthrowing the Protest Paradigm? How The New York Times, Global Voices and Twitter Covered the Egyptian Revolution. *International Journal of Communication*, *5*(0), 16.

Hasyim, M. (2019). Linguistic Functions of Emoji in Social Media Communication. *Opción*, *35*(24), 558–574.

Haupt, M. R., Jinich-Diamant, A., Li, J., Nali, M., & Mackey, T. K. (2021). Characterizing twitter user topics and communication network dynamics of the “Liberate” movement during COVID-19 using unsupervised machine learning and social network analysis. *Online Social Networks and Media*, *21*, 100114. https://doi.org/10.1016/j.osnem.2020.100114

Heemsbergen, L. J., & Lindgren, S. (2014). The power of precision air strikes and social media feeds in the 2012 Israel–Hamas conflict: ‘Targeting transparency.’ *Australian Journal of International Affairs*, *68*(5), 569–591. https://doi.org/10.1080/10357718.2014.922526

Herman, E. S., & Chomsky, N. (2002). *Manufacturing consent: The political economy of the mass media*. Pantheon.

Hermann, R., & Bona, G. (Eds.). (1996). *1848-1849: A szabadságharc és forradalom története*. Videopont.

Hoffman, B. L., Colditz, J. B., Shensa, A., Wolynn, R., Taneja, S. B., Felter, E. M., Wolynn, T., & Sidani, J. E. (2021). #DoctorsSpeakUp: Lessons learned from a pro-vaccine Twitter event. *Vaccine*, *39*(19), 2684–2691. https://doi.org/10.1016/j.vaccine.2021.03.061

Honnibal, M., & Montani, I. (2017). *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*.

Hornsey, M. J., Blackwood, L., Louis, W., Fielding, K., Mavor, K., Morton, T., O'Brien, A., Paasonen, K.-E., Smith, J., & White, K. M. (2006). Why Do People Engage in Collective Action? Revisiting the Role of Perceived Effectiveness. *Journal of Applied Social Psychology*, *36*(7), 1701–1722. https://doi.org/10.1111/j.0021-9029.2006.00077.x

Howard, P. N., Duffy, A., Freelon, D., Hussain, M. M., Mari, W., & Maziad, M. (2011). *Opening Closed Regimes: What Was the Role of Social Media During the Arab Spring?* (SSRN Scholarly Paper ID 2595096; Issue ID 2595096). Social Science Research Network. https://doi.org/10.2139/ssrn.2595096

Huang, Y., Fei, T., Kwan, M.-P., Kang, Y., Li, J., Li, Y., Li, X., & Bian, M. (2020). GIS-Based Emotional Computing: A Review of Quantitative Approaches to Measure the Emotion Layer of Human–Environment Relationships. *ISPRS International Journal of Geo-Information*, *9*(9), 551. https://doi.org/10.3390/ijgi9090551

Hunnicutt, T. (2020, September 25). *Biden slams Trump for refusing to speak out about Belarus 'dictator' Lukashenko*. Global News. https://globalnews.ca/news/7361001/belarus-biden-trump-dictator/

Hunt, S. A., & Benford, R. D. (1994). Identity Talk in the Peace and Justice Movement. *Journal of Contemporary Ethnography*, *22*(4), 488–517. https://doi.org/10.1177/089124194022004004

Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), 216–225.

IEA. (2020, April). *Belarus energy profile – Analysis*. IEA. https://www.iea.org/reports/belarus-energy-profile

Ilie, L. (2017, February 10). 'We see you'—Romanian activists become potent political force. *Reuters*. https://www.reuters.com/article/us-romania-corruption-activists-idUSKBN15P1KA

Ince, J., Rojas, F., & Davis, C. A. (2017). The social media response to Black Lives Matter: How Twitter users interact with Black Lives Matter through hashtag use. *Ethnic and Racial Studies*, *40*(11), 1814–1830. https://doi.org/10.1080/01419870.2017.1334931

Jamal, A. A., Keohane, R. O., Romney, D., & Tingley, D. (2015). Anti-Americanism and Anti-Interventionism in Arabic Twitter Discourses. *Perspectives on Politics*, *13*(1), 55–73. https://doi.org/10.1017/S1537592714003132

Jasper, J. M. (2011). Emotions and Social Movements: Twenty Years of Theory and Research. *Annual Review of Sociology*, *37*(1), 285–303. https://doi.org/10.1146/annurev-soc-081309-150015

Jasper, J. M., & Poulsen, J. D. (1995). Recruiting Strangers and Friends: Moral Shocks and Social Networks in Animal Rights and Anti-Nuclear Protests. *Social Problems*, *42*(4), 493–512. https://doi.org/10.2307/3097043

Jenkins, J. C. (1983). Resource Mobilization Theory and the Study of Social Movements. *Annual Review of Sociology*, *9*(1), 527–553. https://doi.org/10.1146/annurev.so.09.080183.002523

Jones, L. K. (2020). #BlackLivesMatter: An Analysis of the Movement as Social Drama. *Humanity & Society*, *44*(1), 92–110. https://doi.org/10.1177/0160597619832049

Jungherr, A., & Jürgens, P. (2014). Through a Glass, Darkly: Tactical Support and Symbolic Association in Twitter Messages Commenting on Stuttgart 21. *Social Science Computer Review*, *32*(1), 74–89. https://doi.org/10.1177/0894439313500022

Jungherr, A., Jürgens, P., & Schoen, H. (2012). Why the Pirate Party Won the German Election of 2009 or The Trouble With Predictions: A Response to Tumasjan, A., Sprenger, T. O., Sander, P. G., & Welpe, I. M. "Predicting Elections With Twitter: What 140 Characters Reveal About Political Sentiment." *Social Science Computer Review*, *30*(2), 229–234. https://doi.org/10.1177/0894439311404119

Kaplan, S. L. (1985). The Paris Bread Riot of 1725. *French Historical Studies*, *14*(1), 23. https://doi.org/10.2307/286413

Kardara, M., Papadakis, G., Papaoikonomou, A., Tserpes, K., & Varvarigou, T. (2015). Large-scale evaluation framework for local influence theories in Twitter. *Information Processing & Management*, *51*(1), 226–252. https://doi.org/10.1016/j.ipm.2014.06.002

Keck, M. E., & Sikkink, K. (1998). *Activists beyond Borders: Advocacy Networks in International Politics*. Cornell University.

Kenney, K., Smith, K., Moriarty, S., & Barbatsis, G. (Eds.). (2005). *Handbook of visual communication: Theory, methods, and media*. LEA.

Kharroub, T., & Bas, O. (2016). Social media and protests: An examination of Twitter images of the 2011 Egyptian revolution. *New Media & Society*, *18*(9), 1973–1992. https://doi.org/10.1177/1461444815571914

Khondker, H. H. (2011). Role of the New Media in the Arab Spring. *Globalizations*, *8*(5), 675–679. https://doi.org/10.1080/14747731.2011.621287

Kim, Y., Hsu, S.-H., & de Zúñiga, H. G. (2013). Influence of Social Media Use on Discussion Network Heterogeneity and Civic Engagement: The Moderating Role of Personality Traits. *Journal of Communication*, *63*(3), 498–516. https://doi.org/10.1111/jcom.12034

Klandermans, B. (1984). Mobilization and Participation: Social-Psychological Expansisons of Resource Mobilization Theory. *American Sociological Review*, *49*(5), 583–600. https://doi.org/10.2307/2095417

Klandermans, B., van der Toorn, J., & van Stekelenburg, J. (2008). Embeddedness and Identity: How Immigrants Turn Grievances into Action. *American Sociological Review*, *73*(6), 992–1012. https://doi.org/10.1177/000312240807300606

Kornhauser, W. (2017). *The Politics of Mass Society*. Routledge.

Korolov, R., Lu, D., Wang, J., Zhou, G., Bonial, C., Voss, C., Kaplan, L., Wallace, W., Han, J., & Ji, H. (2016). On predicting social unrest using social media. *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 89–95. https://doi.org/10.1109/ASONAM.2016.7752218

Krasimirov, A. (2018, May 12). Bulgarian government to seek mandate for talks with investors over nuclear plant. *Reuters*. https://www.reuters.com/article/us-bulgaria-nuclear-belene/bulgarian-government-to-seek-mandate-for-talks-with-investors-over-nuclear-plant-idUSKCN1ID0TE

Krosigk, E. von, & Sighele, S. (2008). *Psychologie des Auflaufs und der Massenverbrechen*. VDM Verlag Dr. Müller.

Kwon, K. H., Wang, H., Raymond, R., & Xu, W. W. (2015). A spatiotemporal model of Twitter information diffusion: An example of Egyptian revolution 2011. *Proceedings of the 2015 International Conference on Social Media & Society - SMSociety '15*, 1–7. https://doi.org/10.1145/2789187.2789205

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting & Electronic Media*, *40*(4), 460–477. https://doi.org/10.1080/08838159609364369

Larmer, B. (2018, November 1). What 52,000 Percent Inflation Can Do to a Country. *The New York Times*. https://www.nytimes.com/2018/11/01/magazine/venezuela-inflation-economics.html

Latour, B., & LaTour, C. de S. de l'Innovation B. (2005). *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University.

Le Bon, G., & Merton, R. K. (1960). *The crowd: A study of the popular-mind*. Viking Press.

Leach, C. W., Iyer, A., & Pedersen, A. (2006). Anger and Guilt About Ingroup Advantage Explain the Willingness for Political Action. *Personality and Social Psychology Bulletin*, *32*(9), 1232–1245. https://doi.org/10.1177/0146167206289729

LeFebvre, R. K., & Armstrong, C. (2018). Grievance-based social movement mobilization in the #Ferguson Twitter storm. *New Media & Society*, *20*(1), 8–28. https://doi.org/10.1177/1461444816644697

Lindgren, S., & Lundström, R. (2011). Pirate culture and hacktivist mobilization: The cultural and social protocols of #WikiLeaks on Twitter. *New Media & Society*, *13*(6), 999–1018. https://doi.org/10.1177/1461444811414833

Linfield, S. (2011). *The Cruel Radiance: Photography and Political Violence*. University of Chicago Press.

Lipsky, M. (1968). Protest as a Political Resource. *The American Political Science Review*, *62*(4), 1144–1158. https://doi.org/10.2307/1953909

Lipsky, M. (1969). *Protest in city politics: Rent strikes, housing, and the power of the poor*. Rand McNally.

Loria, S. (2018). *Textblob* (0.15, 2.) [Computer software].

Lotan, G., Graeff, E., Ananny, M., Gaffney, D., Pearce, I., & Boyd, D. (2011). The Arab Spring| The Revolutions Were Tweeted: Information Flows during the 2011 Tunisian and Egyptian Revolutions. *International Journal of Communication*, *5*(0), 31.

Lupia, A., & Sin, G. (2003). Which public goods are endangered? How evolving communication technologies affect "The Logic of Collective Action." *Public Choice*, *117*, 315–331.

Lyebyedyev, Y., & Makhortykh, M. (2018a). #Euromaidan: Quantitative Analysis of Multilingual Framing 2013–2014 Ukrainian Protests on Twitter. *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, 276–280. https://doi.org/10.1109/DSMP.2018.8478462

Lyebyedyev, Y., & Makhortykh, M. (2018b). #Euromaidan: Quantitative Analysis of Multilingual Framing 2013–2014 Ukrainian Protests on Twitter. *2018 IEEE Second International Conference on Data Stream Mining Processing (DSMP)*, 276–280. https://doi.org/10.1109/DSMP.2018.8478462

Lynch, M. (2011). After Egypt: The Limits and Promise of Online Challenges to the Authoritarian Arab State. *Perspectives on Politics*, *9*(2), 301–310.

Lysenko, V. V., & Desouza, K. C. (2012). Moldova's internet revolution: Analyzing the role of technologies in various phases of the confrontation. *Technological Forecasting and Social Change*, *79*(2), 341–361. https://doi.org/10.1016/j.techfore.2011.05.009

Macafee, T., & De Simone, J. J. (2012). Killing the Bill Online? Pathways to Young People's Protest Engagement via Social Media. *Cyberpsychology, Behavior, and Social Networking*, *15*(11), 579–584. https://doi.org/10.1089/cyber.2012.0153

Maireder, A., & Schwarzenegger, C. (2012). A Movement of Connected Individuals. *Information, Communication & Society*, *15*(2), 171–195. https://doi.org/10.1080/1369118X.2011.589908

Makhovsky, A. (2020, August 12). Thousands stage flower protest in Belarus as EU weighs sanctions. *Reuters*. https://www.reuters.com/article/us-belarus-election-idUSKCN2581OQ

Margetts, H., John, P., Hale, S., & Taha, Y. (2015). *Political Turbulence: How Social Media Shape Collective Action*. Princeton University.

Marshall, G. (1998). *A Dictionary of Sociology*. Oxford University Press.

McAdam, D. (1986). Recruitment to High-Risk Activism: The Case of Freedom Summer. *American Journal of Sociology*, *92*(1), 64–90.

McAdam, D., Tarrow, S. G., & Tilly, C. (2001). *Dynamics of contention*. Cambridge University.

McCarthy, J. D., & Zald, M. N. (1973). *The Trend of Social Movements in America: Professionalization and Resource Mobilization*. General Learning Press.

McCarthy, J. D., & Zald, M. N. (1977). Resource Mobilization and Social Movements: A Partial Theory. *American Journal of Sociology*, *82*(6), 1212–1241.

Melucci, A. (1980). The new social movements: A theoretical approach. *Social Science Information*, *19*(2), 199–226. https://doi.org/10.1177/053901848001900201

Melucci, A. (1985). The Symbolic Challenge of Contemporary Movements. *Social Research*, *52*(4), 789–816.

Melucci, A. (1988). Getting involved: Identity and mobilization in social movements. *International Social Movement Research*, *1*(26), 329–348.

Melucci, A. (1989). *Nomads of the Present: Social Movements and Individual Needs in Contemporary Society*. Hutchinson Radius.

Melucci, A. (1995). The Process of Collective Identity. In H. Johnston & B. Klandermans (Eds.), *Social Movements and Culture* (pp. 41–63). UCL Press.

Melucci, A. (1996). *Challenging Codes. Collective Action in the Information Age*. Cambridge University.

Melucci, A. (2009). *Challenging Codes: Collective Action in the Information Age*. Cambridge University. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=4637514

Meraz, S., & Papacharissi, Z. (2013). Networked Gatekeeping and Networked Framing on #Egypt. *The International Journal of Press/Politics*, *18*(2), 138–166. https://doi.org/10.1177/1940161212474472

Mercea, D. (2013). Probing the Implications of Facebook Use for the Organizational Form of Social Movement Organizations. *Information, Communication & Society*, *16*(8), 1306–1327. https://doi.org/10.1080/1369118X.2013.770050

Mercea, D., Burean, T., & Proteasa, V. (2020). Student Participation and Public Facebook Communication: Exploring the Demand and Supply of Political Information in the Romanian #rezist Demonstrations. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3684089

Meyer, D. S., & Minkoff, D. C. (2004). Conceptualizing Political Opportunity. *Social Forces*, *82*(4), 1457–1492.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *ArXiv:1301.3781 [Cs]*. http://arxiv.org/abs/1301.3781

Milan, S. (2013). The Italian anomaly. *Index on Censorship*, *42*(1), 12–15. https://doi.org/10.1177/0306422013477017

Mirror. (2011, August 25). London riots: More than 2,000 people arrested over disorder. *Mirror*. https://www.mirror.co.uk/news/uk-news/london-riots-more-than-2000-people-185548

Mitchell, L., Frank, M. R., Harris, K. D., Dodds, P. S., & Danforth, C. M. (2013). The Geography of Happiness: Connecting Twitter Sentiment and Expression, Demographics, and Objective Characteristics of Place. *PLOS ONE*, *8*(5), e64417. https://doi.org/10.1371/journal.pone.0064417

Molina, B. (2017, October 26). *Twitter overcounted active users since 2014, shares surge on profit hopes*. USA TODAY. https://www.usatoday.com/story/tech/news/2017/10/26/twitter-overcounted-active-users-since-2014-shares-surge/801968001/

Morales, A., Losada González, J. C., & Benito, R. (2012). Users Structure and Behavior on an Online Social Network During a Political Protest. *Physica A: Statistical Mechanics and Its Applications*, *391*, 5244–5253. https://doi.org/10.1016/j.physa.2012.05.015

Moran, M., & Waddington, D. (2016). Violence and Looting on the Streets of London: The English Riots of 2011. In M. Moran & D. Waddington, *Riots* (pp. 115–140). Palgrave Macmillan UK. https://doi.org/10.1057/978-1-137-57131-1_6

Morris, A. D., & Mueller, C. M. (1992). *Frontiers in Social Movement Theory*. Yale University.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the Sample Good Enough? Comparing Data from Twitter's Streaming API with Twitter's Firehose. *ArXiv:1306.5204 [Physics]*. http://arxiv.org/abs/1306.5204

Moss, M. L. (1987). Telecommunications, World Cities, and Urban Policy. *Urban Studies*, *24*(6), 534–546.

MTI. (2018, March 10). *Orbán: Events in Slovakia 'bear Soros's fingerprints.'* Daily News Hungary. https://dailynewshungary.com/orban-events-slovakia-bear-soross-fingerprints/

Mueller, M. L. (2010). *Networks and States: The Global Politics of Internet Governance*. MIT.

Muller, C. L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., & Leigh, R. R. (2015). Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *International Journal of Climatology*, *35*(11), 3185–3203. https://doi.org/10.1002/joc.4210

Muller, C. l., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., & Leigh, R. r. (2015). Crowdsourcing for climate and atmospheric sciences: Current status and future potential. *International Journal of Climatology*, *35*(11), 3185–3203. https://doi.org/10.1002/joc.4210

Munger, K., Bonneau, R., Nagler, J., & Tucker, J. A. (2019). Elites Tweet to Get Feet Off the Streets: Measuring Regime Social Media Strategies During Protest. *Political Science Research and Methods*, *7*(04), 815–834. https://doi.org/10.1017/psrm.2018.3

Nacher, A. (2021). #BlackProtest from the web to the streets and back: Feminist digital activism in Poland and narrative potential of the hashtag. *European Journal of Women's Studies*, *28*(2), 260–273. https://doi.org/10.1177/1350506820976900

Nez, H. (2021). What has become of the Indignados? The biographical consequences of participation in the 15M movement in Madrid (2011–19). *Social Movement Studies*, 1–20. https://doi.org/10.1080/14742837.2021.1977113

Nunns, A., & Idle, N. (Eds.). (2011). *Tweets from Tahrir: Egypt's Revolution as it Unfolded, in the Words of the People Who Made it*. OR Books.

OccupyWallSt. (2011). *About Us*. http://occupywallst.org/about/

Oh, O., Eom, C., & Rao, H. R. (2015). **Research Note** —Role of Social Media in Social Change: An Analysis of Collective Sense Making During the 2011 Egypt Revolution.

Information Systems Research, *26*(1), 210–223. https://doi.org/10.1287/isre.2015.0565

Oh, O., Kwon, K., Agrawal, M., & Rao, R. (2011). *Choice of Information: A Study of Twitter News Sharing during the 2009 Israel-Gaza Conflict. 2*.

Olesen, T. (2013a). *Injustice symbols: On the political-cultural outcomes of social movements*.

Olesen, T. (2013b). “We are all Khaled Said”: Visual Injustice Symbols in the Egyptian Revolution, 2010–2011. In N. Doerr, A. Mattoni, & S. Teune (Eds.), *Research in Social Movements, Conflicts and Change* (Vol. 35, pp. 3–25). Emerald Group Publishing Limited. https://doi.org/10.1108/S0163-786X(2013)0000035005; http://web.archive.org/web/20200412195633/https://www.emerald.com/insight/content/doi/10.1108/S0163-786X(2013)0000035005/full/html

Oliinyk, A., & Kuzio, T. (2021). The Euromaidan Revolution, Reforms and Decommunisation in Ukraine. *Europe-Asia Studies*, *73*(5), 807–836. https://doi.org/10.1080/09668136.2020.1862060

Olson, M. (1971). *The Logic of Collective Action*. Harvard University Press.

Pak, A., & Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*. LREC 2010, Valletta, Malta. http://www.lrec-conf.org/proceedings/lrec2010/pdf/385_Paper.pdf

Panagiotopoulos, P., Bigdeli, A. Z., & Sams, S. (2014). Citizen–government collaboration on social media: The case of Twitter in the 2011 riots in England. *Government Information Quarterly*, *31*(3), 349–357. https://doi.org/10.1016/j.giq.2013.10.014

Papacharissi, Z., & Oliveira, M. de F. (2012). Affective News and Networked Publics: The Rhythms of News Storytelling on #Egypt. *Journal of Communication*, *62*(2), 266–282. https://doi.org/10.1111/j.1460-2466.2012.01630.x

Park, S. J., Lim, Y. S., & Park, H. W. (2015). Comparing Twitter and YouTube networks in information diffusion: The case of the “Occupy Wall Street” movement. *Technological Forecasting and Social Change*, *95*, 208–217. https://doi.org/10.1016/j.techfore.2015.02.003

Paul, I., Khattar, A., Kumaraguru, P., Gupta, M., & Chopra, S. (2019). Elites Tweet? Characterizing the Twitter Verified User Network. *2019 IEEE 35th International Conference on Data Engineering Workshops (ICDEW)*, 278–285. https://doi.org/10.1109/ICDEW.2019.00006

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Peña-López, I., Congosto, M., & Aragón, P. (2014). Spanish Indignados and the evolution of the 15M movement on Twitter: Towards networked para-institutions. *Journal of Spanish Cultural Studies*, *15*(1–2), 189–216. https://doi.org/10.1080/14636204.2014.931678

Penney, J., & Dadas, C. (2014). (Re)Tweeting in the service of protest: Digital composition and circulation in the Occupy Wall Street movement. *New Media & Society*, *16*(1), 74–90. https://doi.org/10.1177/1461444813479593

Poell, T. (2014). Social media and the transformation of activist communication: Exploring the social media ecology of the 2010 Toronto G20 protests. *Information, Communication & Society*, *17*(6), 716–731. https://doi.org/10.1080/1369118X.2013.812674

Poell, T., & Borra, E. (2012). Twitter, YouTube, and Flickr as platforms of alternative journalism: The social media account of the 2010 Toronto G20 protests. *Journalism*, *13*(6), 695–713. https://doi.org/10.1177/1464884911431533

Procter, R., Vis, F., & Voss, A. (2013). Reading the riots on Twitter: Methodological innovation for the analysis of big data. *International Journal of Social Research Methodology*, *16*(3), 197–214. https://doi.org/10.1080/13645579.2013.774172

Qi, H., Jiang, H., Bu, W., Zhang, C., & Shim, K. J. (2019). Tracking Political Events in Social Media: A Case Study of Hong Kong Protests. *2019 IEEE International Conference on Big Data (Big Data)*, 6192–6194. https://doi.org/10.1109/BigData47090.2019.9006462

Räbiger, S., & Spiliopoulou, M. (2015). A framework for validating the merit of properties that predict the influence of a twitter user. *Expert Systems with Applications*, *42*(5), 2824–2834. https://doi.org/10.1016/j.eswa.2014.11.006

Rankin, J., & Leroux, M. (2021, November 21). EU could fund gas project linked to man charged over Maltese journalist's murder. *The Guardian*. https://www.theguardian.com/world/2021/nov/21/eu-could-fund-gas-project-linked-to-man-charged-in-maltese-journalist-murder-daphne-caruana-galizia

Redlands, C. E. S. R. I. (2.8). *ArcGIS Desktop Pro*.

Rehurek, R., & Sojka, P. (2011). Gensim—Python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, *3*(2), 2.

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *ArXiv:1908.10084 [Cs]*. http://arxiv.org/abs/1908.10084

Resch, B., Usländer, F., & Havas, C. (2018). Combining machine-learning topic models and spatiotemporal analysis of social media data for disaster footprint and damage assessment. *Cartography and Geographic Information Science*, *45*(4), 362–376. https://doi.org/10.1080/15230406.2017.1356242

Reuters. (2012a, September 11). *Palestinian ministers meet on protests*. The Irish Times. https://www.irishtimes.com/news/palestinian-ministers-meet-on-protests-1.735954

Reuters. (2012b, November 27). Belarus President Lukashenko in his own words. *Reuters*. https://www.reuters.com/article/us-belarus-lukashenko-extracts-idUSBRE8AQ0V520121127

RND/dpa. (2020, September 30). *"Einfach weggeprügelt": Vor zehn Jahren eskalierte der Stuttgart-21-Protest*. https://www.rnd.de/politik/einfach-weggeprugelt-vor-zehn-jahren-eskalierte-der-stuttgart-21-protest-SF4G5QNS7LD4PRJ4ZQT6VUN3AM.html

Roberts, E., & Stan, C. (2017, 0 7). *Romania protests continue over plans to revive corruption bill*. CNN. https://www.cnn.com/2017/02/06/europe/romania-protests-update/index.html

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408. https://doi.org/10.1145/2684822.2685324

Rodríguez, A. (2020, September 10). *The crisis in Eastern Mediterranean will be the focus of the Med7 summit*. Atalayar. https://atalayar.com/en/content/crisis-eastern-mediterranean-will-be-focus-med7-summit

Röhn, T. (2018, March 10). #AllforJan: Slowakei: Revolte auf Samtpfoten. *Dei Welt*. https://www.welt.de/politik/ausland/article174418602/AllforJan-Slowakei-Revolte-auf-Samtpfoten.html

Ryan, C., & Gamson, W. A. (2006). The Art of Reframing Political Debates. *Contexts*, *5*(1), 13–18. https://doi.org/10.1525/ctx.2006.5.1.13

Ryan, Y. (2011, 0 26). How Tunisia's revolution began. *Al Jazeera*. https://www.aljazeera.com/indepth/features/2011/01/2011126121815985483.html

Schattschneider, E. E. (1960). *The semisovereign people: A realist's view of democracy in America*. Dryden.

Schlegel, A. (1995). My lai: 'We lie, they die': Or, a small history of an 'atrocious' photograph [1]. *Third Text*, *9*(31), 47–66. https://doi.org/10.1080/09528829508576544

Segerberg, A., & Bennett, W. L. (2011). Social Media and the Organization of Collective Action: Using Twitter to Explore the Ecologies of Two Climate Change Protests. *The Communication Review*, *14*(3), 197–215. https://doi.org/10.1080/10714421.2011.597250

Segesten, A. D., & Bossetta, M. (2017). A typology of political participation online: How citizens used Twitter to mobilize during the 2015 British general elections. *Information, Communication & Society*, *20*(11), 1625–1643. https://doi.org/10.1080/1369118X.2016.1252413

Severo, M., & Zuolo, E. (2012). Egyptian e-diaspora: Migrant websites without a network? *Social Science Information*, *51*(4), 521–533. https://doi.org/10.1177/0539018412456772

Shah, N. (n.d.). *emot: Emoji and Emoticons detection package for Python* (3.1) [Computer software]. Retrieved March 26, 2022, from https://github.com/NeelShah18/emo

Shirky, C. (2009). *Here comes everybody: How change happens when people come together*. Penguin. https://archive.org/details/herecomeseverybo0000shir_g7o3

Shirky, C. (2011). The political power of social media: Technology, the public sphere, and political change. *Foreign Affairs*, *90*(1). https://www.semanticscholar.org/paper/The-political-power-of-social-media%3A-Technology%2C-Shirky/73f29492478bb28c422d32f075f688a903addec1

Shoemaker, P. J. (1996). Hardwired for News: Using Biological and Cultural Evolution to Explain the Surveillance Function. *Journal of Communication*, *46*(3), 32–47. https://doi.org/10.1111/j.1460-2466.1996.tb01487.x

Simons, L. M. (1987). *Worth Dying for*. William Morrow & Co.

Sinpeng, A. (2021). Hashtag activism: Social media and the #FreeYouth protests in Thailand. *Critical Asian Studies*, *53*(2), 192–205. https://doi.org/10.1080/14672715.2021.1882866

Smelser. (1963). *Theory of Collective Behavior*. Simon & Schuster.

Snow, D. A., & Bedford, R. D. (1988). Ideology, Frame Resonance, and Participant Mobilization. *International Social Movement Research*, *1*, 197–218.

Snow, D. A., Rochford, E. B., Worden, S. K., & Benford, R. D. (1986). Frame Alignment Processes, Micromobilization, and Movement Participation. *American Sociological Review*, *51*(4), 464–481. https://doi.org/10.2307/2095581

Sontag, S. (1978). *On Photography*. Farrar, Straus and Giroux.

Sotiropoulos, G. (2017). Staging Democracy: The Aganaktismenoi of Greece and the Squares Movement(s). *Contention*, *5*(1). https://doi.org/10.3167/cont.2017.050106

Stage, C. (2013). The online crowd: A contradiction in terms? On the potentials of Gustave Le Bon's crowd psychology in an analysis of affective blogging. *Distinktion: Journal of Social Theory*, *14*(2), 211–226. https://doi.org/10.1080/1600910X.2013.773261

Steiger, E., Resch, B., & Zipf, A. (2016). Exploration of spatiotemporal and semantic clusters of Twitter data using unsupervised neural networks. *International Journal of Geographical Information Science*, *30*(9), 1694–1716. https://doi.org/10.1080/13658816.2015.1099658

Stratfor. (2011, February 1). *Update on the Size of Protests in Cairo*. Stratfor. https://worldview.stratfor.com/article/article/update-size-protests-cairo

Tang, T. Y., & Cheng, M. W. T. (2022). The politicization of everyday life: Understanding the impact of the 2019 Anti-Extradition Law Amendment Bill Protests on pro-democracy protesters' political participation in Hong Kong. *Critical Asian Studies*, *54*(1), 128–148. https://doi.org/10.1080/14672715.2022.2027257

Tarrow, S. (1994). *Power in Movement: Social Movements, Collective Action and Politics*. Cambridge University.

Than, K. (2017, July 11). Hungary's anti-Soros posters recall "Europe's darkest hours": Soros' spokesman. *Reuters*. https://www.reuters.com/article/us-hungary-soros-idUSKBN19W0XU

The Canadian Press. (2020). *$16.5M settlement in class-action lawsuit over mass arrests at 2010 G20 summit—Toronto | Globalnews.ca*. Global News. https://globalnews.ca/news/7281119/lawsuit-2010-g20-summit-toronto-settled/

The Gurardian. (2010, October 1). Protesters clash with German police over Stuttgart 21 rail project. *The Guardian*. http://www.theguardian.com/world/gallery/2010/oct/01/protest-germany-stuttgart-21

Theocharis, Y. (2013). The Wealth of (Occupation) Networks? Communication Patterns and Information Distribution in a Twitter Protest Network. *Journal of Information Technology & Politics*, *10*(1), 35–56. https://doi.org/10.1080/19331681.2012.701106

Theocharis, Y., Lowe, W., van Deth, J. W., & García-Albacete, G. (2015). Using Twitter to mobilize protest action: Online mobilization patterns and action repertoires in the Occupy Wall Street, Indignados, and Aganaktismenoi movements. *Information, Communication & Society*, *18*(2), 202–220. https://doi.org/10.1080/1369118X.2014.948035

Thorpe, H., & Ahmad, N. (2015). Youth, action sports and political agency in the Middle East: Lessons from a grassroots parkour group in Gaza. *International Review for the Sociology of Sport*, *50*(6), 678–704. https://doi.org/10.1177/1012690213490521

Tiernan, P. (2014). A study of the use of Twitter by students for lecture engagement and discussion. *Education and Information Technologies*, *19*(4), 673–690. https://doi.org/10.1007/s10639-012-9246-4

Tillery, A. B. (2019). What Kind of Movement is Black Lives Matter? The View from Twitter. *The Journal of Race, Ethnicity, and Politics*, *4*(2), 297–323. https://doi.org/10.1017/rep.2019.17

Tilly, C. (1978). *From Mobilization to Revolution*. Addison-Wesley.

Tilly, C. (1995). *European revolutions, 1492-1992* (Pbk. ed). Blackwell.

Tonkin, E., Pfeiffer, H., & Tourte, G. (2012). Twitter, information sharing and the London riots? *Bulletin of the American Society for Information Science and Technology*, *38*. https://doi.org/10.1002/bult.2012.1720380212

Tremayne, M. (2014). Anatomy of Protest in the Digital Era: A Network Analysis of Twitter and Occupy Wall Street. *Social Movement Studies*, *13*(1), 110–126. https://doi.org/10.1080/14742837.2013.830969

Tsolova, T. (2013, June 14). Bulgarians protests over media magnate as security chief. *Reuters*. https://www.reuters.com/article/us-bulgaria-government/bulgarians-protests-over-media-magnate-as-security-chief-idUSBRE95D0ML20130614

Tsolova, T., & Krasimirov, A. (2013, July 24). Bulgarians stage new protest rally after siege of parliament. *Reuters*. https://www.reuters.com/article/us-bulgaria-protests-blockade-idUSBRE96N15020130724

Tufekci, Z., & Wilson, C. (2012). Social Media and the Decision to Participate in Political Protest: Observations From Tahrir Square. *Journal of Communication*, *62*(2), 363–379. https://doi.org/10.1111/j.1460-2466.2012.01629.x

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (n.d.). *Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment*. 8.

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape. *Social Science Computer Review*, *29*(4), 402–418. https://doi.org/10.1177/0894439310386557

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2012). Where There is a Sea There are Pirates: Response to Jungherr, Jürgens, and Schoen. *Social Science Computer Review*, *30*(2), 235–239. https://doi.org/10.1177/0894439311404123

Twitter Blog. (2012, March 21). *Twitter turns six*. https://blog.twitter.com/en_us/a/2012/twitter-turns-six

Twitter Inc. (2018). *Twitter API Documentation*. https://developer.twitter.com/en/docs

Valkenburg, P. M., Peter, J., & Walther, J. B. (2016). Media Effects: Theory and Research. *Annual Review of Psychology*, *67*(1), 315–338. https://doi.org/10.1146/annurev-psych-122414-033608

van der Zee, B. (2009). Twitter Triumphs. *Index on Censorship*, *38*(4), 97–102. https://doi.org/10.1080/03064220903392570

van Zomeren, M., Leach, C. W., & Spears, R. (2012). Protesters as "Passionate Economists": A Dynamic Dual Pathway Model of Approach Coping With Collective Disadvantage. *Personality and Social Psychology Review*, *16*(2), 180–199. https://doi.org/10.1177/1088868311430835

van Zomeren, M., Spears, R., Fischer, A. H., & Leach, C. W. (2004). Put Your Money Where Your Mouth Is! Explaining Collective Action Tendencies Through Group-Based Anger and Group Efficacy. *Journal of Personality and Social Psychology*, *87*(5), 649–664. https://doi.org/10.1037/0022-3514.87.5.649

Varnali, K., & Gorgulu, V. (2015). A social influence perspective on expressive political participation in Twitter: The case of #OccupyGezi. *Information, Communication & Society*, *18*(1), 1–16. https://doi.org/10.1080/1369118X.2014.923480

Vatikiotis, P. J. (1997). *The Middle East: From the end of empire to the end of the Cold War*. Routledge.

Veenstra, A. S., Iyer, N., Hossain, M. D., & Park, J. (2014). Time, place, technology: Twitter as an information source in the Wisconsin labor protests. *Computers in Human Behavior*, *31*, 65–72. https://doi.org/10.1016/j.chb.2013.10.011

Vetterkind, R. (2021, February 7). 10 years later: Wisconsin's Act 10 has produced labor savings, but at a cost. *Wisconsin State Journal*. https://madison.com/news/local/govt-and-politics/10-years-later-wisconsins-act-10-has-produced-labor-savings-but-at-a-cost/article_04022e81-82ba-5c23-88f9-25070c031f7c.html

Vicari, S. (2013). Public reasoning around social contention: A case study of Twitter use in the Italian mobilization for global change. *Current Sociology*, *61*(4), 474–490. https://doi.org/10.1177/0011392113479747

Vigil-Hayes, M., Duarte, M., Parkhurst, N. D., & Belding, E. (2017). #Indigenous: Tracking the Connective Actions of Native American Advocates on Twitter. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1387–1399. https://doi.org/10.1145/2998181.2998194

Wang, C.-J., Wang, P.-P., & Zhu, J. J. H. (2013). Discussing Occupy Wall Street on Twitter: Longitudinal Network Analysis of Equality, Emotion, and Stability of Public Discussion. *Cyberpsychology, Behavior, and Social Networking*, *16*(9), 679–685. https://doi.org/10.1089/cyber.2012.0409

Wang, W., Li, Z., Wang, Y., & Chen, F. (2013). Indexing cognitive workload based on pupillary response under luminance and emotional changes. *Proceedings of the 2013 International Conference on Intelligent User Interfaces*, 247–256. https://doi.org/10.1145/2449396.2449428

Wang, Z., Ye, X., & Tsou, M.-H. (2016). Spatial, temporal, and content analysis of Twitter for wildfire hazards. *Natural Hazards*, *83*(1), 523–540. https://doi.org/10.1007/s11069-016-2329-6

Wankel, C. (2009). Management education using social media. *Organization Management Journal*, *6*(4), 251–262. https://doi.org/10.1057/omj.2009.34

Webber, J. R. (2010). Venezuela under Chávez: The Prospects and Limitations of Twenty-First Century Socialism, 1999-2009. *Socialist Studies/Études Socialistes*, *6*(1). https://doi.org/10.18740/S47W2R

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), xiii–xxiii.

Westley, B. H., & MacLean, M. S. (1957). A Conceptual Model for Communications Research. *Journalism Quarterly*, *34*(1), 31–38. https://doi.org/10.1177/107769905703400103

Wikipedia. (2022). George Soros—Wikipedia. In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=George_Soros&oldid=1077913625

Wilkins, D. J., Livingstone, A. G., & Levine, M. (2019). Whose tweets? The rhetorical functions of social media use in developing the Black Lives Matter movement. *British Journal of Social Psychology*, *58*(4), 786–805. https://doi.org/10.1111/bjso.12318

Williams, S. A., Terras, M. M., & Warwick, C. (2013). What do people study when they study Twitter? Classifying Twitter related academic papers. *Journal of Documentation*, *69*(3), 384–410. https://doi.org/10.1108/JD-03-2012-0027

Wojcieszak, M., & Smith, B. (2014). Will politics be tweeted? New media use by Iranian youth in 2011. *New Media & Society*, *16*(1), 91–109. https://doi.org/10.1177/1461444813479594

Yaqub, U., Sharma, N., Pabreja, R., Chun, S. A., Atluri, V., & Vaidya, J. (2018). Analysis and visualization of subjectivity and polarity of Twitter location data. *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, 1–10. https://doi.org/10.1145/3209281.3209313

Ye & Wei. (2019). A Multi-Dimensional Analysis of El Niño on Twitter: Spatial, Social, Temporal, and Semantic Perspectives. *ISPRS International Journal of Geo-Information*, *8*(10), 436. https://doi.org/10.3390/ijgi8100436

Zeitzoff, T. (2011). Using Social Media to Measure Conflict Dynamics: An Application to the 2008–2009 Gaza Conflict. *Journal of Conflict Resolution*, *55*(6), 938–969. https://doi.org/10.1177/0022002711408014

Zelinska, O. (2017). Ukrainian Euromaidan protest: Dynamics, causes, and aftermath. *Sociology Compass*, *11*(9), e12502. https://doi.org/10.1111/soc4.12502

Zhuo, X., Wellman, B., & Yu, J. (2011, July 1). Egypt: The First Internet Revolt? *Peace Magazine*, 6.

Zimmer, M., & Proferes, N. J. (2014). A topology of Twitter research: Disciplines, methods, and ethics. *Aslib Journal of Information Management*, *66*(3), 250–261. https://doi.org/10.1108/AJIM-09-2013-0083