cogitatio

Media and Communication (ISSN: 2183–2439)
2023, Volume 11, Issue 1, Pages 293–305
https://doi.org/10.17645/mac.v11i1.6058

Article

The Multilingual Twitter Discourse on Vaccination in Germany During the Covid-19 Pandemic

Hannah Schmid-Petri [1,*], Moritz Bürger [1], Stephan Schlögl [1], Mara Schwind [1], Jelena Mitrović [2,3], and Ramona Kühn [2]

[1] Faculty of Arts and Humanities, University of Passau, Germany
[2] Faculty of Computer Science and Mathematics, University of Passau, Germany
[3] Institute for Artificial Intelligence Research and Development of Serbia, Serbia

* Corresponding author (hannah.schmid-petri@uni-passau.de)

Submitted: 30 July 2022 | Accepted: 9 January 2023 | Published: 27 March 2023

Abstract
There is evidence that specific segments of the population were hit particularly hard by the Covid-19 pandemic (e.g., people with a migration background). In this context, the impact and role played by online platforms in facilitating the integration or fragmentation of public debates and social groups is a recurring topic of discussion. This is where our study ties in, we ask: How is the topic of vaccination discussed and evaluated in different language communities in Germany on Twitter during the Covid-19 pandemic? We collected all tweets in German, Russian, Turkish, and Polish (i.e., the largest migrant groups in Germany) in March 2021 that included the most important keywords related to Covid-19 vaccination. All users were automatically geocoded. The data was limited to tweets from Germany. Our results show that the multilingual debate on Covid-19 vaccination in Germany does not have many structural connections. However, in terms of actors, arguments, and positions towards Covid-19 vaccination, the discussion in the different language communities is similar. This indicates that there is a parallelism of the debates but no social-discursive integration.

Keywords
content analysis; Covid-19; multilingual communities; Twitter; vaccination debate

Issue
This article is part of the issue "Science Communication in the Digital Age: New Actors, Environments, and Practices" edited by Julia Metag (University of Münster), Florian Wintterlin (University of Münster), and Kira Klinger (University of Münster).

1. Introduction

There is evidence that the population with a migration background in Germany was hit particularly hard by the Covid-19 pandemic, and these groups were less willing to be vaccinated (Robert Koch-Institut, 2021). Various reasons were blamed for this, such as cramped living situations or unfavourable working conditions. Another problem was certainly the fact that, at least at the beginning of the crisis, a lot of information or counselling services were only available in German, which meant that not all population groups were reached.

Social media, especially Twitter, are widely used to engage and discuss the issue of vaccination (e.g., Keim-Malpass et al., 2017; Massey et al., 2016). Thus, social media can reach different societal groups with health-related information. But parallel to the great advantages that online media environments potentially offer for the dialogue between different groups in society, the possibilities of online communication may also lead to a fragmentation of public discourse, meaning that online conversations take place in different homogenous groups which are isolated from one another (Dahlberg, 2007). Also, regarding the process of integration, online

social networks can be used to maintain relationships with the heritage culture, as well as to bond to the new culture and build up social capital (e.g., Kim et al., 2011; Park & Gerrits, 2021). Additionally, a discursive fragmentation regarding most contemporary challenges (e.g., climate change or the Covid-19 pandemic) would be particularly severe; they affect society as a whole, and due to their complexity and possibly huge negative impact, they can only be solved if all societal groups are involved. Up to now, only a few studies have looked at online exchanges between different linguistic communities. The existing research does not provide a uniform picture: While some studies show a fragmentation, for example, between Farsi and English blogs (Kelly & Etling, 2008), others indicate that there is at least some degree of mutual reference and dialogue across language barriers in online social networks or that social media can help to overcome cultural differences (Eleta & Golbeck, 2012; Etling et al., 2014; Hale, 2012, 2014).

Against this background, the overall research question of our study is: How integrated or fragmented is the debate on Twitter across different multilingual communities located in Germany?

To answer our research question, we use the Twitter debate about Covid-19 vaccination as an example and include in our analysis tweets in German, Polish, Turkish, and Russian by Twitter users located in Germany in March 2021. We are particularly interested in how the debate on Twitter originated in Germany is structured, i.e., if the debate breaks into different language communities or concentrates on various positions concerning Covid-19 vaccination regardless of language.

By studying multilingual communication, we contribute to existing research in two main ways. First, our article theoretically combines and integrates research dealing with fragmentation and work regarding the role of media use in the process of social integration. Second, existing research regarding the analysis of public discourses mostly ignores modern societies' cultural and linguistic diversity and seldom analyses processes of fragmentation (or the potential of integration) across different language communities living in the same country. Thus, our study's results help provide deeper insights into how heterogeneous publics form and interact.

2. The Multilingual Discussion About Covid-19 Vaccination on Twitter

Online social networking sites offer new ways for different types of actors to relate to each other—a crucial characteristic of social networking sites like Twitter is their networked character, in which actors form the nodes of a network that can be connected through various types of relations (e.g., follower structures, retweets/forwards, hashtags, or mentions in distinct posts). Thus, each actor using social networking sites individually selects to whom or to which debate or argument they want to relate. The effects of these individual choices are a matter of an ongoing (scientific) debate in which the underlying question always refers to possible fragmentation or integration processes and, thus, the potential (or failure) of the public sphere to integrate different actors and viewpoints (Dahlberg, 2007).

Mostly, fragmentation is understood as a result of a sorting process in which people connect based on common homogeneous characteristics (e.g., ideological or political standpoints; Häussler, 2018), resulting in opposing groups that are segregated from each other.

With regard to our research question, two different drivers of fragmentation (or integration) could be identified: (a) the cultural background and the language in which Twitter users decide to write their tweets, as well as (b) the positions and arguments towards Covid-19 vaccination or regarding certain vaccines (independent from a specific language). Thus, different constellations are possible: First, discourses in the various languages may be detached from each other with varying degrees of internal conflict or consonance. This would point to segregation, i.e., actors have no relations to actors representing another culture but only connections to actors representing their own culture and tweet only in "their" language (Mittelstädt & Odag, 2015). Second, the different publics are connected by opinionative alliances that form around the positions towards the issue at stake with varying degrees of heterogeneity concerning the respective languages. Diverse communities in which actors are linked to each other across different languages could be interpreted as multiple inclusion and, thus, integration. Third, marginalisation would be indicated by isolates in the network, i.e., nodes that have no connections to any other part of the network (Mittelstädt & Odag, 2015).

Only a few studies have looked at the influence language or geographic region has on the structure of online networks. In general, digital media are particularly popular among people with a migration background (Gattringer et al., 2022). In a survey of young migrants (N = 475) from North Rhine-Westphalia (a federal state in Germany), WhatsApp and YouTube were the most used social networking sites; 21% indicated that they use Twitter regularly. But studies also show that language, as well as geography, play a large role in structuring hyperlink networks or follower relationships on Twitter (Hale, 2012, 2014; Herring et al., 2007; Kulshrestha et al., 2012; Takhteyev et al., 2012). Additionally, Hale (2014) found that multilingual users on Twitter were more active than those who tweeted in only one language. Thus, these multilingual users form important bridges between different language communities. Based on this, our first research question is:

> RQ1: How is the multilingual debate about Covid-19 vaccination in Germany on Twitter structured?

Regarding the discussion about vaccination on Twitter, studies show several interesting results. First, the majority of actors tweeting about Covid-19-related vaccination

consist of laypeople sharing their own experiences and opinions; (health or scientific) professionals only seem to play a minor role (Herrera-Peco et al., 2021; Lentzen et al., 2022). Second, the overall sentiment regarding Covid-19 vaccinations seems to be predominantly positive. However, the share of tweets with negative sentiments and mentioning vaccine opposition increased over the course of the pandemic (Bonnevie et al., 2021; Bustos et al., 2022; Hussain et al., 2021; Kwok et al., 2021). Third, regarding the discussed topics, people were particularly focused on vaccine development during the early stages of the pandemic. Pro-vaccine tweets shared their hopes for a timely introduction of a successful vaccine and later praised the relatively fast development, while vaccine-hesitant tweets expressed concern about a (perceived) lack of thorough clinical trials and therefore reduced vaccine safety (Jiang et al., 2021; Liew & Lee, 2021; Lyu et al., 2021; Thelwall et al., 2021). (Perceived) health effects were another widely discussed topic. Tweets with negative vaccine sentiment emphasised (potential) harmful side effects of the Covid-19 vaccines and tried to discourage people from getting vaccinated (Griffith et al., 2021; Liew & Lee, 2021; Muric et al., 2021; Thelwall et al., 2021). However, some people simply documented how they felt after being vaccinated (e.g., mentioning mild side effects) while still supporting and encouraging the uptake of Covid-19 vaccines (Lentzen et al., 2022). Fourth, since the beginning of the pandemic, a number of conspiracy theories have evolved that are disseminated through social networks such as Twitter, e.g. that vaccines will be used to control and monitor the public by implanting microchips into people while vaccinating them (Germani & Biller-Andorno, 2021; Muric et al., 2021). Fifth, specific vaccines were evaluated differently in the discussion on Twitter: BioNTech and Moderna were mostly mentioned in a positive context, whereas AstraZeneca was perceived quite negatively (Jemielniak & Krempovych, 2021; Lyu et al., 2021; Malagoli et al., 2021; Marcec & Likic, 2021).

Existing studies mostly focus on US or English tweets (see Jiang et al., 2021; Kwok et al., 2021; Liew & Lee, 2021; Lyu et al., 2021). However, a few studies focus on non-English discussions, for instance, by analysing tweets from a number of Spanish-speaking countries (e.g., Herrera-Peco et al., 2021). As far as we know, there are even fewer studies looking at different communities within individual countries; notable exceptions are a few works analysing the English Covid-19 vaccination discourse on Twitter (Guntuku et al., 2021; Thelwall et al., 2021). They found that different communities (primarily within the US) focus on varying aspects of the topic, e.g., Black communities debate issues of trust in the healthcare system. Following this, our second and third research questions are:

RQ2: What types of actors participate in the multilingual discourse about Covid-19 vaccination in Germany on Twitter, and which arguments do they use?

RQ3: How similar are the discussions in the different language communities?

3. Methods and Measurement

3.1. Time Frame and Collection of Tweets

We analyse the Twitter discourse on vaccination in Germany during the Covid-19 pandemic. Germany is a country in which a large share of inhabitants has a migration background (approximately 26.7% in 2020; Statistisches Bundesamt, 2022). The German Federal Statistical Office classifies a person as having a migration background if they or at least one parent does not have German citizenship by birth. In 2020, the largest share of people with a migration background (N = 21.9 million) had ties to Turkey at 12.6%, followed by Poland at 9.4% and Russia at 5.6% (Statistisches Bundesamt, 2022). Hence, we included German, Turkish, Polish, and Russian tweets in our sample.

We collected all German, Turkish, Polish, and Russian tweets between March 1st, 2021 and March 31st, 2021, that included one or more keywords related to Covid-19 vaccines. Our goal in defining the keywords was to ensure that our query would capture a wide range of tweets but would also be confined to the debate about vaccines and not Covid-19 in general. Based on the finding of DeVerna et al. (2021) that "covid19," in addition to the generic term "vaccine" as well as the names of certain vaccine developers, captured the most data for English tweets in early 2021, we decided on the following approach: Our list of keywords consisted of "covid19," which is universally applicable across languages, in addition to the generic terms "vaccine" and "vaccination" and the names of all marketing authorisation holders, developers, and specific vaccines that were approved for use or under rolling review in the EU at the time. A translated search query was used for each of the four languages (see Supplementary File). The data was collected by accessing the Twitter v2 API full archive search capability via the academic research program.

All users were then automatically geocoded based on the (non-mandatory) free-form "location" field in their user profile. For this purpose, we used the python library local-geocode (Müller, 2021), which matches (partial) strings against a database export of location names (such as countries, administrative areas, cities, and towns) from geonames.org. We used an extended database export to cover entries with at least 10,000 inhabitants (see Supplementary File for detailed information on matches/non-matches for each language). Relying on "location" fields to geolocate Twitter users is an established approach in social scientific research (for example, Bruns & Enli, 2018; Bruns et al., 2017; Rauchfleisch et al., 2021; Schweinberger et al., 2021; Vogler et al., 2019). However, some groups of users might be more inclined to disclose their location than others. For example, Baruh et al. (2017) find in their extensive

meta-analysis of research on social networking sites that users with greater concerns regarding their privacy are more reluctant to share personal information (see also Schmidt et al., 2022). Thus, users who disclose information about their location are probably less concerned about privacy issues than the general Twitter population. Nevertheless, we believe our geocoding approach is appropriate for this study because the reported differences do not seem to translate directly into differences between language communities. This expectation is supported by previous studies which do not find empirical evidence that a user's cultural background affects their online behaviour regarding private data (Baruh et al., 2017; Liang et al., 2017).

March 2021 was chosen as the investigation period because of the real-world developments that took place at the time, which indicate that it might have been a period when the Twittersphere in Germany was inclined to discuss the issue of vaccination: (a) The German immunisation campaign was making slow progress compared to other countries, (b) the German BioNtech vaccine was already in use, (c) rolling review for the Russian Sputnik V vaccine started in the EU on 4 March 2021, and (d) experts and citizens criticised the AstraZeneca vaccine due to reported severe side effects.

3.2. Sampling

We limited the sample to those 3,216 users who actively participated in the debate with at least two subject-related tweets. Users were considered multilingual if they used two or more of the four languages under study in the sampled tweets (see Table 1 for an overview).

We then built a network of message-sharing activities. This network is formed by considering the sampled users as nodes and their mentions (@user) and retweets of other nodes in the network as edges. Thus, each undirected edge (e_{ij}) represents the tweets authored by user i in which they mentioned or retweeted user j as well as the tweets authored by user j in which user i is retweeted or mentioned. By focusing on mentions and retweets, we capture dynamic interaction patterns between users as opposed to the more static follower structures. In sum, the network consists of 2,229 nodes connected by 4,150 undirected edges.

For the sampling of tweets which became part of the manual coding, a multistep procedure was used. First, all 3,216 active users were included in the sample. Second, we sampled tweets in different languages differently due to limited resources: All Russian (*n* = 420), Turkish (*n* = 88), and Polish (*n* = 39) tweets by active users were included in the sample. Regarding German tweets, we drew a proportionate stratified sample of 18% of all German tweets (*n* = 2658, see Supplementary File).

3.3. Measurement of Actor Types and Their Positions Regarding Covid-19 Vaccination

We conducted a manual quantitative content analysis. The developed codebook included variables on the level of single accounts (e.g., username, main language, actor type) as well as on the level of tweets (main topic, topic sentiment, mentioned vaccines, vaccine evaluation, sentiment towards Covid-19-related vaccination in general; see Supplementary File). Three trained coders worked on the account-related variables and were supported by two more trained coders for the tweet-related variables (see Supplementary File for details). Reliability scores were sufficient for all variables (see Supplementary File).

4. Results

Our first research question concerned how the multilingual debate about Covid-19 vaccination in Germany on Twitter is structured.

The majority of users tweeting in German were part of the network, while only about 30% were isolates (Table 2). However, the opposite was the case for all other languages: Users tweeting in Russian, Turkish, and Polish were mostly isolates and therefore disconnected from the vaccination-related Twittersphere in Germany. In contrast, however, most multilingual users were, via @mentions or retweets, connected to the discursive network that formed around the discussion of Covid-19 vaccination on Twitter.

Table 1. Languages used by Twitter users under study.

Language used	No. of users (perc.)	Mean no. of tweets (sd)
German	3,103 (96.5%)	4.7 (6.8)
Russian	42 (1.3%)	4.9 (4)
Turkish	20 (0.6%)	2.7 (1.3)
Polish	14 (0.4%)	2.6 (1.2)
Multilingual	37 (1.2%)	12.4 (25.1)
Total	3,216 (100%)	4.7 (7.2)

Notes: Based on active users authoring ≥ 2 tweets during the investigation period, *n* = 3,216; users who have published tweets in more than one of these languages were labelled multilingual and excluded from the "German," "Russian," "Turkish," and/or "Polish" categories.

Table 2. Networked and isolated users across all language communities.

Connectivity	German	Russian	Turkish	Polish	Multilingual
Networked users (perc.)	2,183 (70%)	13 (31%)	4 (20%)	0 (0%)	29 (78%)
Isolated users (perc.)	920 (30%)	29 (69%)	16 (80%)	14 (100%)	8 (22%)
Total	3,103 (100%)	42 (100%)	20 (100%)	14 (100%)	37 (100%)

Looking more closely at the multilingual users, almost all switched between two languages, the most frequent pairing being German and Russian (65%, *n* = 37). Interestingly, the only three verified accounts were official channels of Russian consulates and embassies in Germany, which mostly focussed on mentioning the Sputnik V vaccine in a positive way. Additionally, although only 1.2% of all active users were part of this subgroup, multilingual users were comparatively active, with 12.4 tweets per user on average (the average for the whole sample was 4.7 tweets per user, see Table 1). One individual user, in particular, seems responsible for the high mean number of tweets: "Karina Begun," who published 145 tweets in total. Many postings were retweets from accounts that mostly disseminated favourable statements about the Russian government. It is noteworthy that while "Karina Begun" is predominately tweeting in Russian and only occasionally in German, she is more heavily connected to a specific group of German-speaking users in our sample who focus on the negative aspects of vaccine-related policies in their discussions.

Our second research question dealt with the actors participating in the discourse about Covid-19-related vaccination on Twitter and the arguments used in those discussions. To answer this, we first looked at the actor type distribution across the different languages (Figure 1).

All language communities predominately consisted of private citizens, which we understood as individual actors directly translating their offline persona to Twitter and acting in a personal, unofficial capacity (Figure 1). While another large part of the Turkish language community was made up of journalistic actors, there were considerably fewer actors of this actor type in the other language communities. Furthermore, scientific, political, and healthcare actors only played a small role in our sample. Overall, no major differences can be identified regarding the actor type of active Twitter users in the different language communities.

Next, we looked at topics discussed in the different language communities (Figure 2). Across all language communities, vaccine-related policies dominated the discourse. Tweets were coded to contain this topic if they discussed policies directly or indirectly related to Covid-19 vaccines, including, for instance, the (political) vaccine authorisation process or the government's inoculation campaign. The rest of the German tweets were mostly concerned with vaccine-related (health) effects, i.e., discussing vaccine efficacy, health risks, or benefits. At the same time, vaccine availability (on a societal level in terms of production rate, distribution efficiency, and accessibility), vaccine development (in terms of research plans, clinical trials, development progress reports, or funding plans), and the vaccination process (the actual

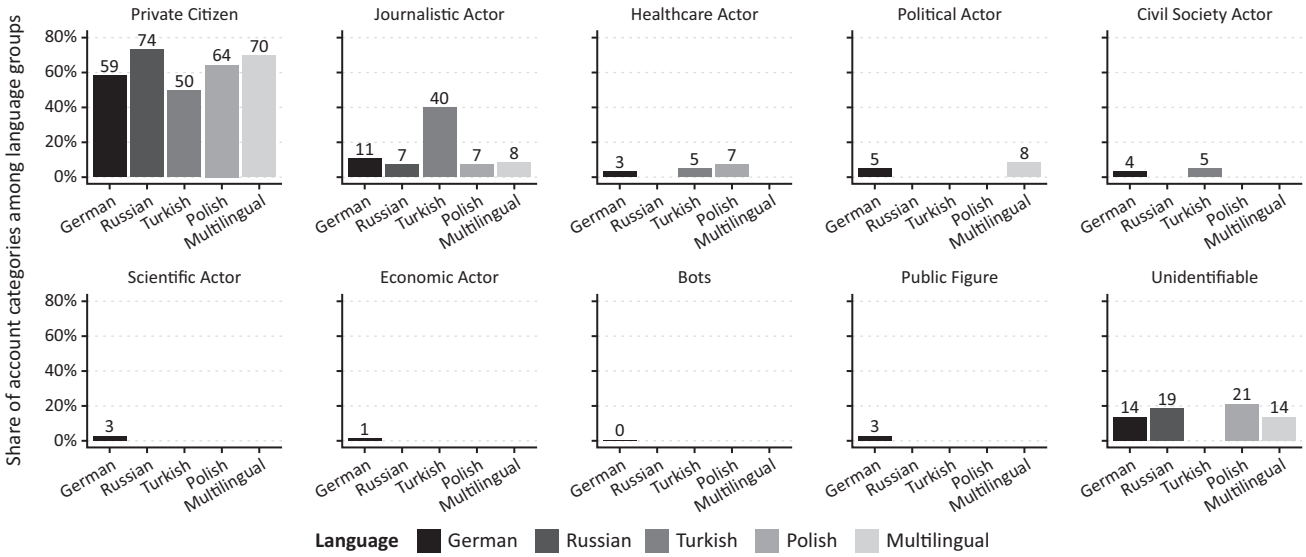

Figure 1. User categories among language groups for accounts with at least two tweets in the sample. Notes: German (*n* = 3,103), Russian (*n* = 42), Turkish (*n* = 20), Polish (*n* = 14), Multilingual (*n* = 37); no statistically significant association (*p* = 0.43).

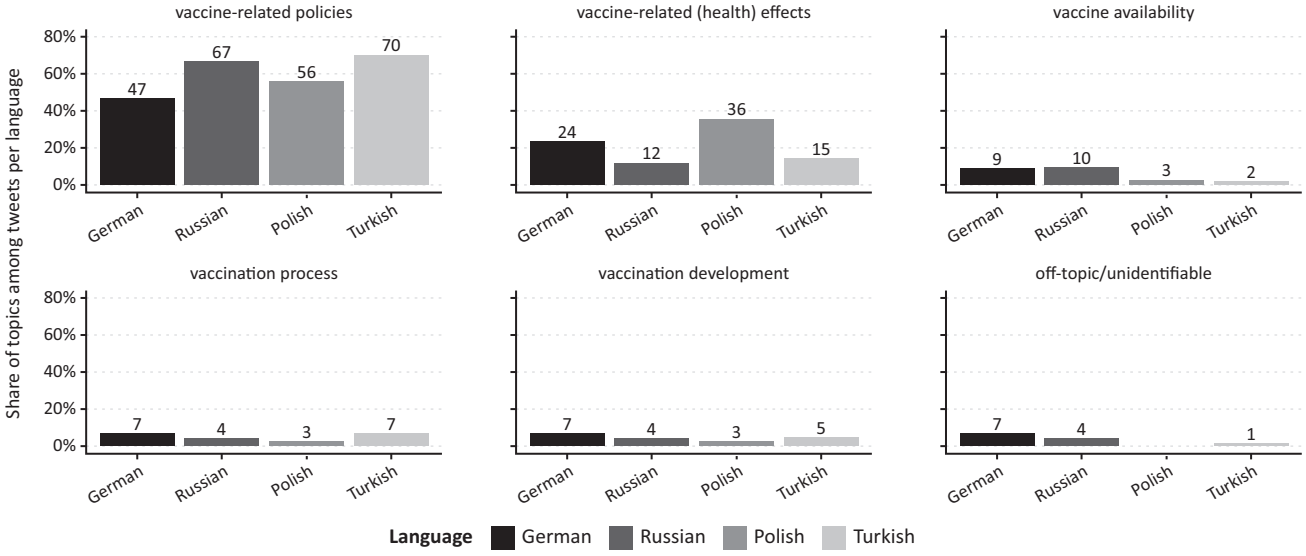

Figure 2. Topic distribution for all coded tweets. Notes: *n* = 3,205, German (*n* = 2,658), Russian (*n* = 420), Turkish (*n* = 88), Polish (*n* = 39); Cramer's *V* = 0.1; *p* < 0.01.

act of getting vaccinated or logistical aspects on an individual or local level) only played a minor role.

This hierarchy of topics was similar for the other languages as well. However, we found some slight differences, i.e., vaccine-related policies being especially important in the Russian and Turkish language communities (see example tweets no. 1–3 in the Supplementary File). In contrast, vaccine-related health effects took up a larger share in the German and Polish language communities in comparison (see example tweets no. 4 and 5 in the Supplementary File).

Regarding sentiments towards Covid-19-related vaccination, the results are similar across all four languages with only slight differences (Figure 3): Most tweets held no subjective opinions about vaccination and discussed the topic neutrally. Interestingly, a positive attitude towards vaccination was the second-most prevalent sentiment, especially in the German and Turkish tweets (see example tweets no. 6 and 7). For instance, users shared encouraging messages about vaccine uptake, demanded better vaccine availability, or discussed the efficacy of vaccination in general. Thus, overall, vaccination is viewed positively on Twitter. In contrast, tweets with negative sentiment contained disparaging messages about Covid-19-related vaccination, i.e., by criticising (perceived) health risks or doubting the efficacy of vaccination.

Sputnik V is the vaccine discussed most frequently in tweets across all four languages (Figure 4). The AstraZeneca vaccine was also discussed frequently, especially in the Polish and German tweets. All other vaccines were mentioned infrequently and did not play a major role in the Covid-19-related vaccination discourse.

Since the discourse seemed to focus on Sputnik V, we were interested in the users' sentiments towards the vaccine (Figure 5). Although the majority of tweets talked about Sputnik V neutrally, the tweets containing a sentiment towards the vaccine were predominately positive and contained supportive messages (see example tweets no. 8–13 in the Supplementary File).

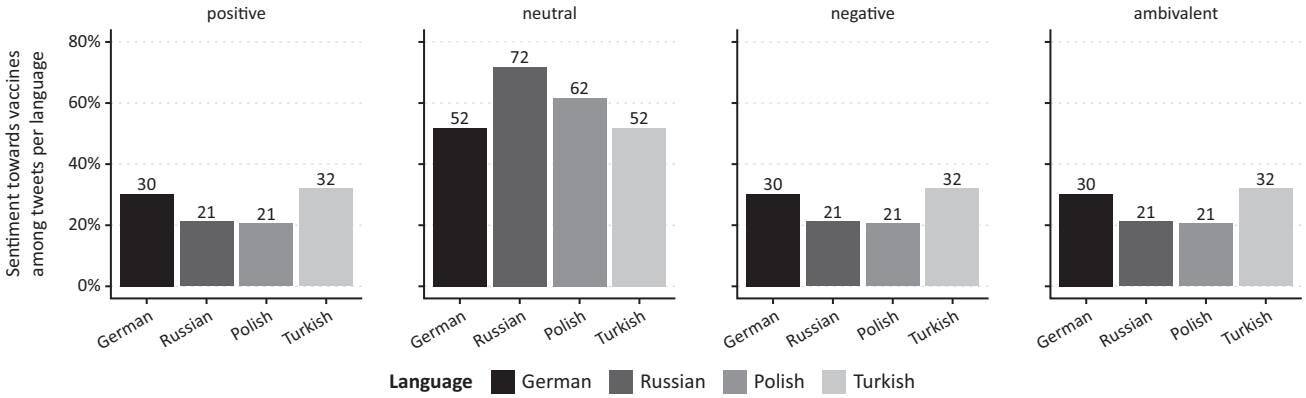

Figure 3. Sentiment distribution for all tweets coded with regard to sentiments about vaccination. Notes: *n* = 3,007, German (*n* = 2,476), Russian (*n* = 405), Turkish (*n* = 87), and Polish (*n* = 39); Cramer's *V* = 0.09; *p* < 0.01.

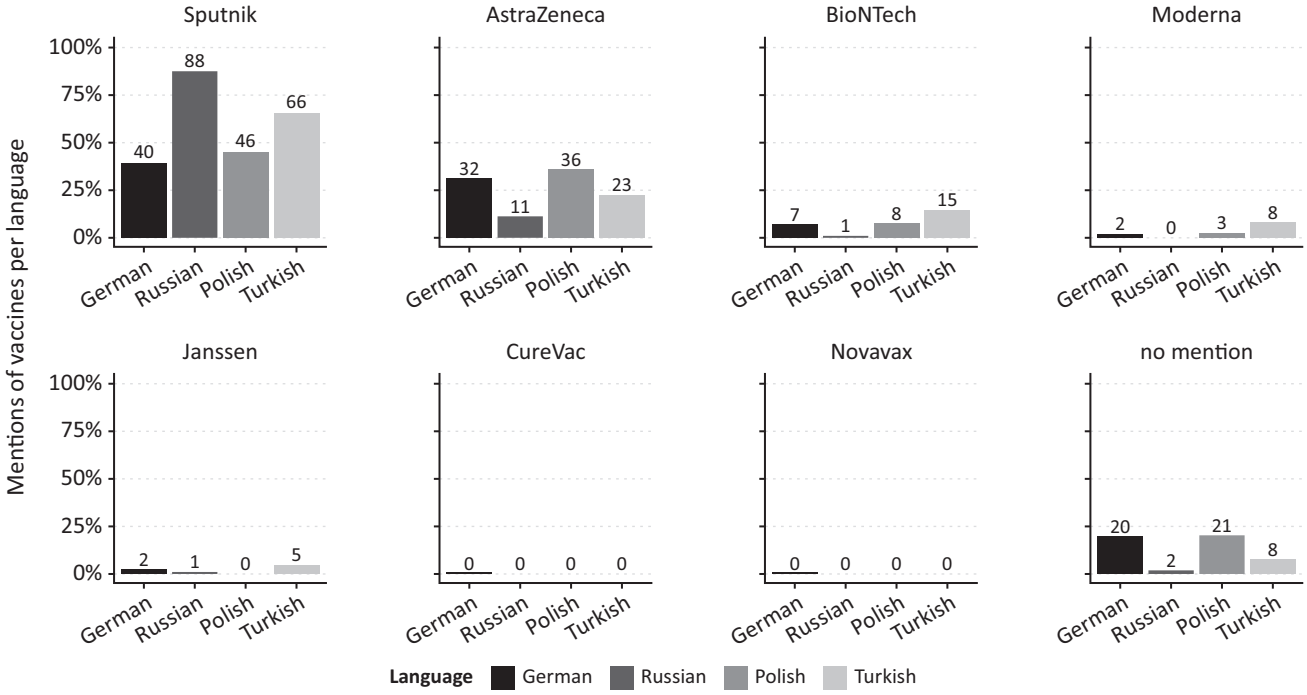

Figure 4. Mentions of specific vaccines in different languages. Notes: The base for each language are all manually coded tweets about Covid-19 vaccination; *n* = 3,205, German (*n* = 2,658), Russian (*n* = 420), Polish (*n* = 39), and Turkish (*n* = 88); Cramer's *V* = 0.19; $p < 0.01$.

Taking a closer look at this, we found that most users authoring those tweets were private individuals using Twitter to criticise the German federal government and the EU and share their personal opinions on various aspects of the Covid-19 pandemic. However, we could not discern a consistent pattern here; the accounts differed in political orientation and their attitude regarding the federal government's Covid-19 policies. The positions were especially divergent on the latter point; many people stated they had had their freedom taken away (see example tweets no. 8 and 9 in the Supplementary File). However, just as many accounts were in favour of containing the Covid-19 pandemic and consequently criticised the government for not going far enough with its measures (see example tweet no. 10 in the Supplementary File). In the next step, we analysed the context in which Sputnik was mentioned. Once again, we found a variety of different lines of argumentation. In most tweets, the users were dissatisfied with either the vaccine availability or the general Covid-19 management of the German federal government. Interestingly, to contextualise and explain those complaints, many users criticised an alleged bias of Germany/the EU against the Sputnik vaccine and Russia itself (see example tweets no. 11 and 12 in the Supplementary File).

In contrast, excluding tweets with a neutral sentiment, the AstraZeneca vaccine was predominately perceived negatively (see Figure 6 and example tweet

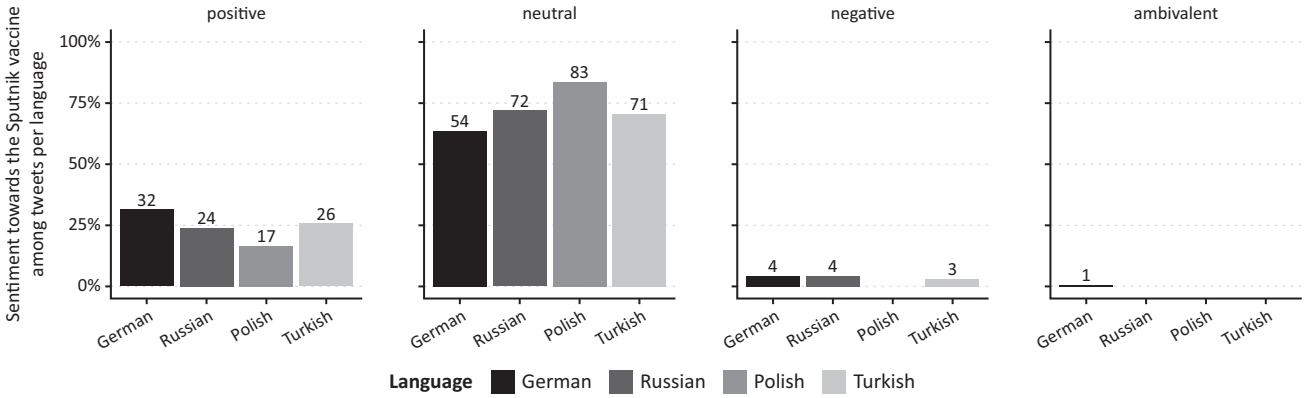

Figure 5. Sentiments towards the Sputnik vaccine. Notes: The base for each language is the number of tweets that mention the Sputnik vaccine; *n* = 1,504, German (*n* = 1,060), Russian (*n* = 368), Polish (*n* = 18), and Turkish (*n* = 58); no statistically significant association ($p = 0.1$).

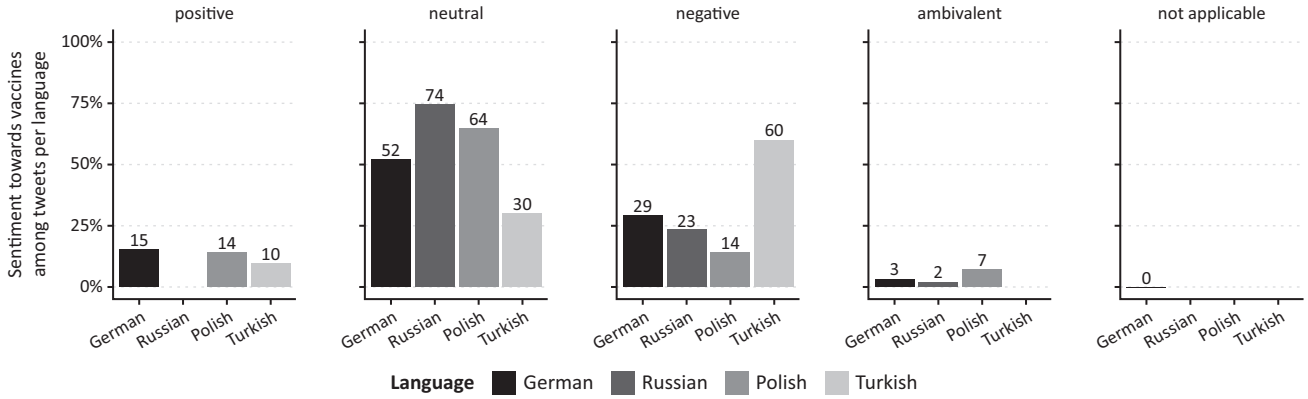

Figure 6. Sentiments towards the AstraZeneca vaccine. Notes: The base for each language is the number of tweets that mention the AstraZeneca vaccine; n = 920, German (n = 839), Russian (n = 47), Polish (n = 14), and Turkish (n = 20); Cramer's V = 0.09; p < 0.01.

no. 8 in the Supplementary File). Only the users tweeting in Polish were undecided since there was an equal number of tweets with positive and negative sentiments.

5. Discussion

In our study, we examined the question of how integrated or fragmented the Covid-19 vaccination debate is on Twitter in Germany.

Overall our results show that in March 2021, comparatively few Twitter users tweeted about Covid-19 vaccinations or revealed their location on Twitter, which were the two main conditions for the sampling of our study. Based on this result, one possible assumption could be that the discussions take place in other social networks, such as Facebook groups or Telegram (see, for example, Peter et al., 2022). In particular, there is little tweeting in other languages by people stating their location in Germany. This may point to the fact that for other cultural groups living in Germany, Twitter does not seem to be relevant for debates on Covid-19 vaccination—at least not for debates in Russian, Polish, or Turkish. Of course, the results could also indicate that people who live in Germany and tweet in other languages are less willing to disclose their location. However, previous research does not suggest that there are cultural differences in the disclosure of geo-location information on Twitter (Liang et al., 2017). Additionally, it is also possible for people with a migration background living in Germany to participate in the vaccination debate on Twitter in German or English.

Regarding our research question, we can state that there are few structural connections between the different language communities. This is especially true for the foreign-language communities, which hardly have any discursive connections via @mentions or retweets to one another or the German community. This is consistent with previous research (e.g., Hale, 2012, 2014)—our data also shows that the Twittersphere is structured along language boundaries. So, when Twitter debates occur in different languages but in the same country, they take place separately and are detached from each other. Multilingual Twitter users seem to act as important bridge actors and liaisons here, again consistent with previous studies (Hale, 2014).

However, regarding the discussed topics and positions towards Covid-19 vaccination and the actor types involved, the debates in the different languages are quite similar, with only a few differences. This means that the debates are not completely independent of each other (i.e., fragmented) in terms of the content discussed.

In all language communities, individuals from civil society dominate the debate—there are hardly any political or scientific actors or organisations from the health sector. Especially during a pandemic, such actors must be present on social networks and actively participate in the debate. Notably, the proportion of journalists is much higher among Turkish Twitter users. One possible (but, of course, speculative) reason could be the repression of media professionals in Turkey, which may have increasingly led Turkish journalists to work from exile. However, it remains unclear why this does not also apply to the Russian community.

Overall, this early Twitter debate shows a positive attitude toward Covid-19 vaccination in all language communities. However, even if most users were neutral or positive about vaccination, the political measures, in particular, were evaluated very negatively, which can be interpreted as criticism of the federal government's Covid-19 management.

The positive evaluation of vaccination in all language communities contradicts the finding that groups with a migration background are less willing to be vaccinated (Robert Koch-Institut, 2021), at least at the beginning of the immunisation campaign. This suggests that foreign-language Twitter users in Germany differ in certain characteristics from other members of the migrant community (e.g., formal education, age, and language skills).

Surprisingly, the discussion on Twitter was particularly positive about Sputnik. In this respect, the debate

on Twitter differs from the discussion of Covid-19 vaccines in other public arenas. Based on the data, it can be assumed that this positive evaluation was due to the fact that other vaccines were (still) in short supply, and thus a better vaccination strategy and a search for alternatives were demanded. Additionally, the positive evaluation of Sputnik was often accompanied by an "East nostalgia" (see example tweet no. 13 in the Supplementary File). Apart from that, the debate about Sputnik shows itself to be a conglomerate of many different attitudes: From the "left" side, a historical entanglement with Russia, which manifests itself in special consideration for Russian interests; from the "right," support for Putin and the strategy of populist elite criticism of German government officials. And last but not least, from the side of the Russian state, the spread of a narrative of Russophobia in the West.

Of course, our study has certain limitations. First, the geocoding approach comes with a few difficulties: In comparison to a random sample of Twitter accounts studied by Hecht et al. (2011), we were not able to locate as many accounts by analysing their location field (66% of users compared to 50.9%). And even if users specified their current location, we could not verify whether this information was correct. The existing research does not indicate any biases related to specific user groups that would substantially tarnish our study's goals (Baruh et al., 2017; Liang et al., 2017; Schmidt et al., 2022). However, digital trace data in general, and self-disclosed user location data in particular, is still afflicted with some uncertainties, such as inferring users' demographics from their social media usage or some users' maverick or even nefarious intentions (for an extensive discussion see Hultquist, 2020). Second, for the coding of the actors, we could only use the information published in their Twitter profiles and the languages of their tweets. This fact also limits our study's validity to some extent since users' migration backgrounds could not be confidently measured. Third, data collected from Twitter is, of course, not representative of the debate among the German population as a whole. Instead, our study considers the role of one platform in a heterogenous public sphere. Fourth, our sample size is rather small and only considers a very limited period of time and a limited number of users and tweets. This applies especially to the foreign language communities. Nevertheless, our data incorporates the entirety of the population tweeting in the non-German languages under study. This shows, as mentioned above, that the group of people tweeting in a foreign language and indicating their location as being in Germany is rather small. To our knowledge, there has not yet been a study that examines the different language communities or the discussion in different languages on Twitter in Germany. Our study could thus be seen as a first attempt to gain a better insight into what is happening on Twitter in this regard and is particularly relevant in light of the debate that various groups with a migration background were hit particularly hard by the pandemic. That said, and with the discussed limitations in mind, our study offers numerous starting points for future research projects: It could be interesting, for example, to combine a survey of Twitter users with a content analysis of their tweets. The survey could, for example, include various demographic variables, the migration background, and the information and usage behaviour of various (social) media. If the users were willing to donate their tweets for scientific purposes, the survey data could be combined with a content analysis of specific tweets, e.g., regarding the discussion about Covid-19 vaccination. This approach would avoid the difficulties of a specific geo-localisation procedure and provide further interesting insights into how people from different (cultural) backgrounds participate in debates on Twitter. Additionally, future research projects could use the manually coded data to establish a "ground truth" for further computational analysis that considers longer investigation periods, larger data sets, or different social networks (e.g., Facebook groups or Telegram). Fifth, we did not explicitly analyse the discussion of conspiracy theories, which would be an interesting line of enquiry concerning potential adverse effects on democracy.

Overall, it can be stated that there are only a few structural connections between Twitter users who tweet in different languages and who indicate in their profile that their location is in Germany. Regarding actors, arguments, and positions towards Covid-19 vaccination, the discussion in the different language communities is similar. This means there is a parallelism of the debates but no social-discursive integration. Thus, given that individuals with migration backgrounds were hit particularly hard by the Covid-19 pandemic, it can be stated that Twitter is not the appropriate channel to protect or engage with these individuals in health debates. Since the debates take place separately in the individual language communities, it seems all the more important to address the communities individually with precisely coordinated communication formats or with local on-site services. However, isolates in our data do not necessarily mean that someone is marginalised. Since we only collected connections to other users in Germany, isolates can also be segregated users who are strongly linked to the community of the country of origin. Additionally, it is important to keep in mind that there is a distinction between online and offline integration based on varying opportunities (Mittelstädt & Odag, 2015).

Acknowledgments

The authors thank the anonymous reviewers for the very helpful and insightful comments. The authors also thank the student assistants Merve Göttl, Veranika Hrusheuskaya, and Isabel Käsbauer for their dedicated help in conducting the manual content analysis. Part of this project (namely the work of Jelena Mitrović and Ramona Kühn) was funded by the German Federal Ministry of Education and Research (BMBF, 01|S20049). The authors are responsible for the content of this publication.

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the authors (unedited).

References

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, *67*(1), 26–53. https://doi.org/10.1111/jcom.12276

Bonnevie, E., Gallegos-Jeffrey, A., Goldbarg, J., Byrd, B., & Smyser, J. (2021). Quantifying the rise of vaccine opposition on Twitter during the Covid-19 pandemic. *Journal of Communication in Healthcare*, *14*(1), 12–19. https://doi.org/10.1080/17538068.2020.1858222

Bruns, A., & Enli, G. (2018). The Norwegian Twittersphere. Structure and dynamics. *Nordicom Review*, *39*(1), 129–148. https://doi.org/10.2478/nor-2018-0006

Bruns, A., Moon, B., Münch, F., & Sadkowsky, T. (2017). The Australian Twittersphere in 2016: Mapping the follower/followee network. *Social Media + Society*, *3*(4), 1–15. https://doi.org/10.1177/2056305117748162

Bustos, V. P., Comer, C. D., Manstein, S. M., Laikhter, E., Shiah, E., Xun, H., Lee, B. T., & Lin, S. J. (2022). Twitter voices: Twitter users' sentiments and emotions about Covid-19 vaccination within the United States. *European Journal of Environment and Public Health*, *6*(1), Article em0096. https://doi.org/10.21601/ejeph/11499

Dahlberg, L. (2007). Rethinking the fragmentation of the cyberpublic: From consensus to contestation. *New Media & Society*, *9*(5), 827–847. https://doi.org/10.1177/1461444807081228

DeVerna, M. R., Pierri, F., Truong, B. T., Bollenbacher, J., Axelrod, D., Loynes, N., Torres-Lugo, C., Yang, K.-C., Menczer, F., & Bryden, J. (2021). CoVaxxy: A collection of English-language Twitter posts about Covid-19 vaccines. *Proceedings of the International AAAI Conference on Web and Social Media*, *15*(1), 992–999. https://doi.org/10.1609/icwsm.v15i1.18122

Eleta, I., & Golbeck, J. (2012). Bridging languages in social networks: How multilingual users of Twitter connect language communities? *Proceedings of the American Society for Information Science and Technology*, *49*(1), 1–4. https://doi.org/10.1002/meet.14504901327

Etling, B., Kelly, J., Faris, R., & Palfrey, J. (2014). Mapping the Arabic blogosphere: Politics, culture, and dissent. In L. Hudson, A. Iskandar, & M. Kirk (Eds.), *Media evolution on the eve of the Arab Spring* (pp. 49–74). Palgrave Macmillan. https://doi.org/10.1057/9781137403155_4

Gattringer, K., Mohr, I., & Rühle, A. (2022). Mediennutzung von Menschen mit internationaler Geschichte: Ergebnisse der Studie ARD/ZDF-Massenkommunikation Trends 2021 [Media usage of people with an international history: Results of the 2021 ARD/ZDF mass communication trends study]. *Media Perspektiven*, *2022*(1), 2–17.

Germani, F., & Biller-Andorno, N. (2021). The anti-vaccination infodemic on social media: A behavioral analysis. *PloS One*, *16*(3), Article e0247642. https://doi.org/10.1371/journal.pone.0247642

Griffith, J., Marani, H., & Monkman, H. (2021). Covid-19 vaccine hesitancy in Canada: Content analysis of tweets using the theoretical domains framework. *Journal of Medical Internet Research*, *23*(4), Article e26874. https://doi.org/10.2196/26874

Guntuku, S. C., Buttenheim, A. M., Sherman, G., & Merchant, R. M. (2021). Twitter discourse reveals geographical and temporal variation in concerns about Covid-19 vaccines in the United States. *Vaccine*, *39*(30), 4034–4038. https://doi.org/10.1016/j.vaccine.2021.06.014

Hale, S. A. (2012). Net increase? Cross-lingual linking in the blogosphere. *Journal of Computer-Mediated Communication*, *17*(2), 135–151. https://doi.org/10.1111/j.1083-6101.2011.01568.x

Hale, S. A. (2014). Global connectivity and multilinguals in the Twitter network. In *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 833–842). Association for Computing Machinery.

Häussler, T. (2018). Heating up the debate? Measuring fragmentation and polarisation in a German climate change hyperlink network. *Social Networks*, *54*, 303–313. https://doi.org/10.1016/j.socnet.2017.10.002

Hecht, B., Hong, L., Suh, B., & Chi, E. H. (2011). Tweets from Justin Bieber's heart: The dynamics of the "location" field in user profiles. In *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 237–246). Association for Computing Machinery.

Herrera-Peco, I., Jiménez-Gómez, B., Romero Magdalena, C. S., Deudero, J. J., García-Puente, M., Benítez De Gracia, E., & Ruiz Núñez, C. (2021). Antivaccine movement and Covid-19 negationism: A content analysis of Spanish-written messages on Twitter. *Vaccines*, *9*(6), Article 656. https://doi.org/10.3390/vaccines9060656

Herring, S. C., Paolillo, J. C., Ramos-Vielba, I., Kouper, I., Wright, E., Stoerger, S., Scheidt, L. A., & Clark, B. (2007). Language networks on LiveJournal. In R. H. Sprague (Ed.), *HICSS '07: Proceedings of the 40th Annual Hawaii International Conference on System Sciences* (p. 79). IEEE.

Hultquist, C. (2020). Representation in geosocial data: Grappling with uncertainty in digital traces of human activity. *International Journal of Humanities and Arts Computing*, *14*(1/2), 218–234. https://doi.org/10.3366/ijhac.2020.0253

Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., & Sheikh, A. (2021). Artificial intelligence-enabled analysis of public attitudes on Facebook and Twitter toward COVID-19 vaccines in the United Kingdom and the United States: Observational study. *Journal of Medical Internet Research*, *23*(4), Article e26627. https://doi.org/10.2196/26627

Jemielniak, D., & Krempovych, Y. (2021). An analysis of AstraZeneca COVID-19 vaccine misinformation and fear mongering on Twitter. *Public Health*, *200*, 4–6. https://doi.org/10.1016/j.puhe.2021.08.019

Jiang, L. C., Chu, T. H., & Sun, M. (2021). Characterisation of vaccine tweets during the early stage of the COVID-19 outbreak in the United States: Topic modeling analysis. *JMIR Infodemiology*, *1*(1), Article e25636. https://doi.org/10.2196/25636

Keim-Malpass, J., Mitchell, E. M., Sun, E., & Kennedy, C. (2017). Using Twitter to understand public perceptions regarding the #HPV vaccine: Opportunities for public health nurses to engage in social marketing. *Public Health Nursing*, *34*(4), 316–323. https://doi.org/10.1111/phn.12318

Kelly, J., & Etling, B. (2008). *Mapping Iran's online public: Politics and culture in the Persian blogosphere*. Berkman Klein Center for Internet and Society. https://cyber.harvard.edu/publications/2008/Mapping_Irans_Online_Public

Kim, Y., Sohn, D., & Choi, S. M. (2011). Cultural difference in motivations for using social network sites: A comparative study of American and Korean college students. *Computers in Human Behavior*, *27*(1), 365–372. https://doi.org/10.1016/j.chb.2010.08.015

Kulshrestha, J., Kooti, F., Nikravesh, A., & Gummadi, K. P. (2012). Geographic dissection of the Twitter network. *Proceedings of the Sixth International AAAI Conference on Web and Social Media*, *6*(1), 202–209. https://doi.org/10.1609/icwsm.v6i1.14280

Kwok, S. W. H., Vadde, S. K., & Wang, G. (2021). Tweet topics and sentiments relating to COVID-19 vaccination among Australian Twitter users: Machine learning analysis. *Journal of Medical Internet Research*, *23*(5), Article e26953. https://doi.org/10.2196/26953

Lentzen, M.-P., Huebenthal, V., Kaiser, R., Kreppel, M., Zoeller, J. E., & Zirk, M. (2022). A retrospective analysis of social media posts pertaining to COVID-19 vaccination side effects. *Vaccine*, *40*(1), 43–51. https://doi.org/10.1016/j.vaccine.2021.11.052

Liang, H., Shen, F., & Fu, K.-w. (2017). Privacy protection and self-disclosure across societies: A study of global Twitter users. *New Media & Society*, *19*(9), 1476–1497. https://doi.org/10.1177/1461444816642210

Liew, T. M., & Lee, C. S. (2021). Examining the utility of social media in COVID-19 vaccination: Unsupervised learning of 672,133 Twitter posts. *JMIR Public Health and Surveillance*, *7*(11), Article e29789. https://doi.org/10.2196/29789

Lyu, J. C., Le Han, E., & Luli, G. K. (2021). Covid-19 vaccine-related discussion on Twitter: Topic modeling and sentiment analysis. *Journal of Medical Internet Research*, *23*(6), Article e24435. https://doi.org/10.2196/24435

Malagoli, L. G., Stancioli, J., Ferreira, C. H. G., Vasconcelos, M., Da Couto Silva, A. P., & Almeida, J. M. (2021). A look into COVID-19 vaccination debate on Twitter. In *13th ACM Web Science Conference 2021* (pp. 225–233). Association for Computing Machinery.

Marcec, R., & Likic, R. (2021). Using Twitter for sentiment analysis towards AstraZeneca/Oxford, Pfizer/BioNTech and Moderna COVID-19 vaccines. *Postgraduate Medical Journal*, *98*(1161), 544–550. https://doi.org/10.1136/postgradmedj-2021-140685

Massey, P. M., Leader, A., Yom-Tov, E., Budenz, A., Fisher, K., & Klassen, A. C. (2016). Applying multiple data collection tools to quantify human papillomavirus vaccine communication on Twitter. *Journal of Medical Internet Research*, *18*(12), Article e318. https://doi.org/10.2196/jmir.6670

Mittelstädt, A., & Odag, Ö. (2015). Social media use and social integration of ethnic minorities in Germany: A new interdisciplinary framework. *Athens Journal of Mass Media and Communications*, *2*(1), 21–32. https://doi.org/10.30958/ajmmc.2.1.2

Müller, M. (2021). *Local-geocode.* GitHub. https://github.com/mar-muel/local-geocode

Muric, G., Wu, Y., & Ferrara, E. (2021). COVID-19 vaccine hesitancy on social media: Building a public Twitter dataset of anti-vaccine content, vaccine misinformation and conspiracies. *JMIR Public Health and Surveillance*, *7*(11), Article e30642. https://doi.org/10.2196/30642

Park, S., & Gerrits, L. (2021). How migrants manifest their transnational identity through online social networks: Comparative findings from a case of Koreans in Germany. *Comparative Migration Studies*, *9*, Article 10. https://doi.org/10.1186/s40878-020-00218-w

Peter, V., Kühn, R., Mitrović, J., Granitzer, M., & Schmid-Petri, H. (2022). Network analysis of German COVID-19 related discussions on Telegram. In P. Rosso, V. Basile, R. Martínez, E. Métais, & F. Meziane (Eds.), *Natural language processing and information systems* (pp. 25–32). Springer.

Rauchfleisch, A., Vogler, D., & Eisenegger, M. (2021). Public sphere in crisis mode: How the COVID-19 pandemic influenced public discourse and user behaviour in the Swiss Twitter-sphere. *Javnost—*

The Public, *28*(2), 129–148. https://doi.org/10.1080/13183222.2021.1923622

Robert Koch-Institut. (2021). *COVID-19 Impfquoten-Monitoring in Deutschland (COVIMO): Report 5—Fokuserhebung Impfquoten* [COVID-19 vaccination rate monitoring in Germany (COVIMO): Report 5—Focus survey vaccination rates]. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/COVIMO_Reports/covimo_studie_bericht_5.pdf?__blob=publicationFile

Schmidt, P., Gordoni, G., Ajzen, I., Beuthner, C., Davidov, E., Silber, H., Steinmetz, H., & Weiß, B. (2022). Twitter users' privacy behavior: A reasoned action approach. *Social Media + Society*, *8*(3), 1–18. https://doi.org/10.1177/20563051221126085

Schweinberger, M., Haugh, M., & Hames, S. (2021). Analysing discourse around COVID-19 in the Australian Twittersphere: A real-time corpus-based analysis. *Big Data & Society*, *8*(1), 1–17. https://doi.org/10.1177/20539517211021437

Statistisches Bundesamt. (2022). *Bevölkerung und Erwerbstätigkeit. Bevölkerung mit Migrationshintergrund: Ergebnisse des Mikrozensus 2020* [Population and employment. Population with a migration background: Results of the microcensus 2020]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Migration-Integration/Publikationen/Downloads-Migration/migrationshintergrund-endergebnisse-2010220207004.pdf?__blob=publicationFile

Takhteyev, Y., Gruzd, A., & Wellman, B. (2012). Geography of Twitter networks. *Social Networks*, *34*(1), 73–81. https://doi.org/10.1016/j.socnet.2011.05.006

Thelwall, M., Kousha, K., & Thelwall, S. (2021). Covid-19 vaccine hesitancy on English-language Twitter. *Profesional de la Información*, *30*(2), Article e300212. https://doi.org/10.3145/epi.2021.mar.12

Vogler, D., Rauchfleisch, A., Eisenegger, M., & Schwaiger, L. (2019). Agenda-Setting auf Twitter—welche Rolle spielen Informationsmedien in der Schweizer Twitter-Sphäre? [Agenda setting on Twitter—What role do news media play in the Swiss Twittersphere?]. In Forschungszentrum Öffentlichkeit und Gesellschaft (Ed.), *Qualität der Medien. Schweiz—Suisse—Svizzera. Jahrbuch 2019* [Quality of the media. Switzerland. Yearbook 2019.] (pp. 47–57). Schwabe. https://doi.org/10.5167/uzh-177417

About the Authors

Hannah Schmid-Petri is a professor of science communication at the University of Passau. Her research is situated at the intersection of political and science communication. She is also a member of the Board of Directors of Bavarian Research Institute for Digital Transformation (bidt). Before her time in Passau, she was a senior assistant at the Institute of Communication and Media Studies at the University of Bern.

Moritz Bürger is a PhD candidate at the University of Passau's Chair of Science Communication. He studied communication science, political science, and peace research. His teaching and research interests include science communication, political communication, and quantitative research methods.

Stephan Schlögl is a PhD student at the University of Passau, Germany. His research and teaching interests include political communication in social media, climate change communication, public sphere theory, and computational social sciences.

Mara Schwind is a research assistant and PhD candidate at the Chair of Science Communication at the University of Passau. In her research, she focuses on digital science communication in social networks as well as other online spaces and is particularly interested in the reciprocal effects of criticism, harassment, and hostility directed at scientists.

Jelena Mitrović (PhD) is a junior research group leader at the Faculty of Computer Science and Mathematics, University of Passau. Her research group CAROLL (Computational Rhetoric in Social Media and Law), funded by the German Ministry of Science and Education (BMBF), focuses on natural language processing topics such as hate speech detection, argumentation mining, and legal text processing.

Ramona Kühn is a research assistant and PhD candidate in CAROLL research group at the University of Passau. After obtaining a master's degree in Computer Science, she now focuses on analysing social media in the context of rhetorical figures. Her goal is to develop algorithms and models to identify and classify such figures. In this process, she created formal domain ontologies for rhetorical figures in German and one combining German, Serbian, and English.