International Journal of Communication 18(2024), 191–218 1932–8036/20240005

Online Toxicity Against Syrians in Turkish Twitter: Analysis and Implications

HALA MULKI♦[1]
The ORSAM Center for Middle Eastern Studies, Turkey

SAMIR ALABDULLAH
Harmoon Center for Contemporary Studies, Turkey

AHMED HALIL
Selçuk University, Turkey

NAWARI AL-ALI
University of Aleppo, Syria

MARIA KYRIAKIDOU
Cardiff University, UK

LUDEK STAVINOHA
University of East Anglia, UK

This study examines the portrayal of Syrians on Turkish Twitter between January and August 2021 through a big data analysis of more than 30,000 tweets. We employ the concept of online toxicity to differentiate between disinformation and hate speech and explore how they are embedded in the negative debates about Syrians on Twitter. Through opinion analysis, the study recognizes disinformation and hate speech patterns within tweets and questions the role they play in boosting anti-Syrian narratives, as well as the main actors behind them in the Turkish Twittersphere. The findings indicate that the

Hala Mulki: halamulki@gmail.com
Samir Alabdullah: samiralabdalh85@yahoo.com
Ahmed Halil: ahmethalil8686@gmail.com
Nawari Al-Ali: nawar_alali@hotmail.com
Maria Kyriakidou: kyriakidoum@cardiff.ac.uk
Ludek Stavinoha: l.stavinoha@uea.ac.uk
Date submitted: 2023-03-19

[1] This research was supported by funding from the Cara Council for At-Risk Academics (Cara), Syria Program Grants 2020, United Kingdom.

discourse about Syrians on Twitter was overwhelmingly negative, with both disinformation and hate speech playing a significant role. Furthermore, a considerable portion of the disinformation tweets could be traced back to opposition political actors, highlighting how negative sentiment on Twitter was not only expressive of generalized public resentment against Syrians but also instrumentalized for political purposes. Overall, this article demonstrates how Twitter contributes to the public debate about Syrians in Turkey, reproducing nationalist narratives and serving political agendas.

Keywords: Twitter, refugees, big data, disinformation, hate speech, online toxicity

The Syrian refugee crisis constitutes a major consequence of the continuous armed conflict in Syria. With Turkey hosting more than 3.5 million Syrian refugees (United Nations High Commissioner for Refugees [UNHCR], 2021), the country's anti-Syrian sentiment has been growing (Karakas, 2021). This is largely the result of politicizing the question of Syrians' status in Turkish society, especially with changes in the political scene. After the 2017 referendum, Turkey witnessed the emergence of two major pre-electoral alliances: The first is the ruling People's Alliance (Cumhur İttifakı), which mainly includes the Justice and Development Party (AKP) and the Nationalist Movement Party (MHP). The second represents the opposition Nation's Alliance (Millet İttifakı), which consists of the Republican People's Party (CHP), Good Party (İyi), the Islamist Felicity Party, Democrat Party, the Democracy and Progress Party, and the Future Party (Secen, Serhun, & Bekir, 2023). Most of these opposition parties, besides the newly emerged Victory Party (Zafer), employed promises to expel Syrians from the country as part of their election campaigns (Farooq, 2021). These arguments have become the focus of many debates across social media, particularly Twitter. Syrian refugees have been misrepresented and have become a target of online toxicity in terms of disinformation, hate speech, and digital racism (Ozduzen, 2020). Beyond psychological harm, such online toxicity has fueled physical and verbal hate crimes against Syrian refugees (Özerim & Tolay, 2021). An example of such an attack was the August 2021 violent riots against Syrians in the Altındağ neighborhood in Ankara ("Turkish Capital," 2021).

Relevant research in media studies has investigated online discourses against refugees, particularly in the aftermath of the 2015 "refugee crisis" at Europe's borders (Erdogan-Ozturk & Isik-Guler, 2020). Most of these studies have focused on the social aspects of digital racism and affiliated it with Turkish populism (Özerim & Tolay, 2021). Others have integrated quantitative analysis of a single hashtag that was trending in a specific context, such as the Taksim Square Protests (2019; Idiz, 2019), and investigated relevant tweets (Ozduzen, 2020). This body of research has highlighted the circulation of hateful discourses and "online toxicity" (Kim, Guess, Nyhan, & Reifler, 2021) against refugees in Turkey and beyond.

Nevertheless, online toxicity has not yet been adequately conceptualized in terms of its different components or the social and political actors instigating it. In this article, we seek to explore the prominent discourses about Syrians on Turkish Twitter, how negative sentiments against them relate to disinformation and hate speech, and the actors behind disinformation and hatred campaigns. Our findings demonstrate that Twitter is not only reflective of a generalized negative sentiment in Turkish society against Syrians but, rather, reproduces and amplifies toxic phenomena, such as disinformation and hate speech. This is motivated by the fact that with more than 18 million active users in Turkey, Twitter has grown popular among Turkish political leaders, as they represent 40% of the top 15 Twitter users (Statista Search Department, 2023). Therefore, we find that online toxicity directed against the Syrian population is often fabricated and affiliated with a specific political agenda related to the debates between the ruling coalition and the opposition. We argue that these hate speech campaigns on social media are strategically instrumentalized by identifiable political actors rooted within particular social and spatial contexts.

Thus, the contribution of our study is twofold. Empirically, we unpack the concept of online toxicity to illustrate the distinction and interplay between online hate speech and disinformation. Methodologically, we employ a novel framework that utilizes a machine learning–based sentiment analysis model proposed by Mulki, Haddad, Ali, & Babaoğlu (2018) and natural language processing (NLP) techniques to collect and investigate more than 30,000 Turkish tweets during the period January–August 2021 and recognize hate speech and disinformation tweets, along with their source geolocations. Overall, this article provides research insights into both discussions on the mediation of forced migration and academic debates on online hate speech and disinformation.

Syrian Refugees on Social Media

The role of social media platforms in migration discourses has been underscored by ambivalence. On one hand, social media has been seen as providing space for alternative representations of the hostility afforded to migrants by mainstream media. While media coverage has been largely negative, if not overtly hostile, against migrants and refugees in different national contexts (Chouliaraki & Stolic, 2017), social media has provided a space for alternative representations, where solidarity for migrants and refugees can be expressed, challenging dominant policies and mainstream coverage (Siapera, 2019). Furthermore, social media platforms have provided refugees with a "voice" (Georgiou, 2018) within the digital media space and opportunities for social support and resilience (Udwan, Leurs, & Alencar, 2020), as well as enabled tracing their migration route patterns even when based on the approximate locations in geotagged tweets (Hübl, Cvetojevic, Hochmair, & Paulus, 2017).

On the other hand, significant research has illustrated that dominant discourses on social media have echoed hostile mainstream media coverage. The hashtag #refugeesnotwelcome, for example, was employed in different European contexts to reproduce a broader rhetoric of exclusion, constructing refugees as outsiders and threatening criminals (Kreis, 2017). Although refugee narratives on social media are far from monolithic, processes of othering of refugees are not only present but even more dehumanizing than mainstream media representations, as hashtags such as #rapefugees indicate.

These broader themes are echoed in the portrayal of Syrians in Turkey. Hate speech against Syrians in Turkish media was already present in 2015 (Az, Gelişli, Barak, & Arslan, 2017) and has been increasing since. Syrians have been negatively stereotyped in terms of their culture, as well as being an economic burden to Turkey, in ways that have contributed to expressions of aggression and violence against them (Alp, 2018). The least negative coverage, according to Şen (2017), can be found in *Hürriyet* and *HaberTürk*, both newspapers supporting the governing AKP, and evidently following its more open policy toward refugees.

In social media, negative stereotypes are further amplified. According to (Aydınlı, 2020), Syrians have been criticized for not fighting for their country and for financially benefiting from Turkey, while also constituting a threat to the local population. Disinformation shared on social media platforms further reiterates the myth of the perceived privileges enjoyed by Syrians. Twitter has become a forum for "digital racism," allowing for the racialization of refugees and the circulation of xenophobia (Ozduzen, 2020). Overall, the diverse rhetorical practices of othering Syrians on Twitter have included their construction as criminals, cowards and traitors of their homelands, invaders of Turkey, possible terrorists, fake refugees, and "lesser" Muslims than Turkish people (Erdogan-Ozturk & Isik-Guler, 2020).

These negative portrayals have served broader populist and nationalist discourses, reaffirming relations of inclusion and exclusion and, ultimately, the national narrative (Aridici, 2022). The hashtag *#ÜlkemdesuriyeliİstemiyoruM,* whose English translation is (#IdontwantSyriansinmycountry), on Twitter has been used to express and reinforce ideas of a strong and exclusionary nationalist identity (Erdogan-Ozturk & Isik-Guler, 2020). Studying Turkish Twitter in the aftermath of Erdogan's statement about giving Syrians citizenship rights in the summer of 2016 (Bozdağ, 2020) found that even the minority of tweets supporting citizenship for Syrians adopted a nationalist frame, either doing so based on the imagination of Islamic brotherhood or emphasizing the importance of security investigations for those considered for citizenship. Similarly, her research on how Syrians were represented during COVID-19 both in mainstream and online news media in Turkey (Yücel, 2021) found that Syrian refugees were largely ignored or "symbolically annihilated" in media coverage of the pandemic, even though they faced the pandemic's greatest socioeconomic consequences. The few news stories that mentioned them focused on Turkey's "generous" policies while criticizing European immigration policies.

These studies illustrate that Twitter discussions are expressive of a general negative sentiment and its instrumentalization within specific nationalist discourses and agendas. Therefore, it is important to study online toxicity not only in terms of hate speeches but also disinformation. Although the concept has been used in various ways and to different ends by communication scholars (Masullo Chen, Muddiman, Wilner, Pariser, & Stroud, 2019), following the definition provided by Kim et al. (2021), we understand online toxicity as the political comments "expressing disrespect for someone by using insulting language, profanity, or name-calling; by engaging in personal attacks; and/or by employing racist, sexist, and xenophobic terms" (p. 924). The multidimensionality of the concept, which includes threats, obscenity, insults, hate, harassment, and "socially disruptive persuasion, such as misinformation and radicalization" (Sheth, Shalin, & Kursuncu, 2021, p. 3), allows accounting for the variety of ways in which negative sentiments and hostile discourses against Syrians are expressed on Turkish Twitter. What we mainly focus on in this article is the interplay between what can be understood as hate speech and instrumental practices of disinformation, both important components of online toxicity. This focus allows us to move beyond depoliticized notions of "uncivil" or "impolite" speech (Masullo Chen et al., 2019) to explore how online toxicity can be purposefully employed as a political tool.

Despite the lack of a universally accepted definition of hate speech (Alkiviadou, 2019), a widely used conceptualization is Nockleby (2000), which describes the concept as "any communication that disparages a person or a group based on some characteristics, such as race, color, ethnicity, gender, sexual orientation, nationality, religion, or other characteristics" (p. 1277). The circulation of hate speech on social media platforms has been a concern for media scholars, especially with far-right discourses in different contexts (e.g., Vidgen & Yasseri, 2020). Calls for a regulatory framework to tackle the ease and speed with which hateful content circulates online (Alkiviadou, 2019) have been met slowly and reluctantly by the adoption of self-regulation practices by different social media platforms, including Twitter. Understanding online toxicity against Syrians as hate speech allows us to detect the type of hateful Twitter comments that can be seen not only as discriminatory and racist but also as inciting violence against them.

The second dimension of online toxicity considered in this article is that of disinformation, namely, the dissemination of deliberately falsified information. There is a "discursive affinity" between hate speech and disinformation, as their tactics and aims often align when the target of spreading falsified information is to instigate and promote negative sentiment against specific groups (Hameleers, van der Meer, & Vliegenthart, 2021). "Hate groups" and ideologues, for example, White supremacists, are among the dominant groups spreading disinformation online (Marwick & Lewis, 2017). At the same time, however, it is important to distinguish between the two concepts, hate speech and disinformation, as hate speech is not always the result of disinformation, and disinformation is not exclusively aimed at spreading hate. Exploring them separately allows us to identify the nuances of Twitter debates and trace the production and circulation of online toxicity to the communicative agency of strategic deceptions deployed by influential political actors. However, we believe that the process of producing and spreading online toxicity can be further investigated when aligned with the geographic context. Such a context is an important aspect to consider when conducting refugee-related studies (Hübl et al., 2017). Consequently, we sought to address the following research questions:

RQ1: What are the prominent discourses being circulated about Syrians on Turkish Twitter?

RQ2: How does the negative stance against Syrians on Twitter relate to disinformation and hate speech?

RQ3: Which actors are behind the online smear campaigns against Syrians?

Methodology

In seeking to identify patterns of online toxicity and disinformation against Syrian refugees in the Turkish Twittersphere, we conducted a big data study. We collected the tweets using the trending hashtags related to Syrians on Turkish Twitter to first identify the negative views within the sample through a sentiment analysis process and then explicate instances of disinformation or hate speech among them, while also considering the associated geolocation information for each tweet category. To these ends, we employed a framework in which a large-scale collection of Turkish tweets was scraped and then subjected to NLP and machine-learning techniques to extract the views in these tweets and identify patterns of disinformation and hate speech.

We mined the tweets during the period January–August 2021 to identify changes over the year. We took the antirefugee riots in Ankara in August 2021 as the end point of our sample. Using Twitter's application programming interface, we harvested tweets along with their associated metadata, such as the username, location, number of followers, number of retweets, etc. This enabled avoiding the limitations of data size, quality, and availability faced by previous studies that used off-the-shelf scraping software (Ozduzen, 2020; Özerim & Tolay, 2021). The collection process relied on top trending hashtags (Table 1) about Syrians because of intensive engagement. This allows for a better capture of the main discourses circulated about Syrians, the most common portrayals of Syrian refugees, and the public stance toward them. We also tracked the most-liked and most-retweeted relevant tweets and scraped the tweet replies posted within their threads. Some of the collected tweets were harvested from the timelines of influential users in Turkey (proponents and opponents of hosting Syrians).

Table 1. The Trending Hashtags Used to Scrape Tweets.

Hashtag (Turkish)	Hashtag (English)
#SuriyelileriAlmayın	*#Don't Host Syrians*
#SuriyelilerinVatanıSuriyedir	*#Syrian's Home is Syria*
#UlkemdeMülteciİstemiyorum	*#I don't Want Refugees in My Country*
#ProvokasyonaGelme	*#Don't trigger Provocation*
#Suriyelileriİstemiyoruz	*#We don't want Syrians*
#KardeşimeDokunma	*#Don't Hurt My Brother*

The raw collected data were composed of 276,093 tweets. Except for the influential users—politicians, celebrities, and other public figures—the author names (usernames) of the tweets were masked throughout the analysis. The raw collected tweets were then subjected to several cleaning and normalization procedures, such as excluding irrelevant tweets (e.g., tweets containing only Afghan-related keywords or hashtags) and objective (opinion-free, i.e., news) tweets. We also applied a stemming technique using a Zemberek stemmer (Akın & Akın, 2007) and reduced Twitter-inherited symbols, punctuation, and stopwords (Mulki et al., 2018). While this step has been ignored in the literature (Assimakopoulos, Baider, & Millar, 2017; Kreis, 2017), we believe tweet cleaning and normalization are crucial to conducting an accurate opinion identification and thus reaching context-relevant conclusions. Consequently, we ended up with a collection of 33,100 Syrian-related, subjective tweets. These tweets were posted by 18,738 unique accounts during the timeframe of our study (January 1–August 31, 2021). These were both from international and Turkish locations, as shown in Table 2.

Table 2. Statistics of the Tweet Collection.

Property	#Tweets/Accounts
Subjective Tweets	33,100
Unique Accounts	18,738
Tweets having geolocations	12,177
Tweets having real geolocations	10,701

For sentiment analysis, unlike similar studies that relied totally on the author's judgment to identify the sentiment in the tweets (Kreis, 2017), we employed a machine learning–based opinion analysis model for Turkish. Mulki et al. (2018) developed this model and configured it to recognize negative and positive opinions embedded in Turkish tweets based on specific combinations of linguistic and stylistic features. It should be noted that while previous research tends to use off-the-shelf sentiment analysis software as it is (Ozduzen, 2020), we opted to customize the model presented in Mulki et al. (2018) in terms of enriching the training data with a manually annotated subset of the studied tweets. The manual annotation for positive/negative tweets was provided by two Turkish native speakers, with an inter-rater agreement of 87%. Additionally, we split the annotated tweet data set into training, validation, and test sets using a stratified sampling technique to ensure that the training, validation, and test sets had approximately the same percentage of samples of each target class as the complete set (Wang, Dai, Shen, & Xuan, 2021). To do so, we took a random 20% of the tweets from each class for the validation and test sets. This enhanced the model's generalization ability and improved evaluation and reliability.

To further mine the negative tweets and detect those containing disinformation, we adopted a hybrid method that combines word-based analysis with syntactic analysis, such that word collocations and specific syntactic patterns are both considered. This is motivated by the fact that truth-tellers and liars have distinctive writing styles. For instance, the authors of disinformation tweets tended to use fewer self-oriented pronouns than other-oriented pronouns, along with other, more sensory-based words (Stahl, 2018). The hybrid method combined the following steps:

- Spotting the most frequent word collocations/associations (pairs/triples) based on a scoring approach (TF-IDF;[2] Mondal, Sahoo, Wang, Mondal, & Rahman, 2022)
- Using the top 1,000 pairs/triples as search queries to filter the tweets that contain them.
- Identifying Syrian-related disinformation content using multiword terms (MWTs) and the Zemberek syntactic parser (Akın & Akın, 2007). Multiword terms are the meaning indicators of a sentence/document (Henry, Cuffy, & McInnes, 2018; Mulki, Ali, Haddad, & Babaoğlu, 2019). In our case, they represent certain syntactic patterns (MWTs) that can refer to allegations (topics) about Syrians. In such syntactic patterns, proper nouns like "Suriyeli/Suriyeliler" *(Syrian/Syrians)* may collocate with nouns such as "Maaş" *(salary)* or proper nouns such as "Esad" *(Bashar Assad)*, or verbs like "geziyorlar" *(traveling)*, and verbs with negations like "Mazlum değil" *(not oppressed)* such word collocations are inspired by the social/political context of Syrian refugees in Turkey.

To detect hate speech, we follow Waseem, Davidson, Warmsley, & Weber (2017), who argue that hate speech discourses are mainly composed of abusive language, which makes hate speech detection a subtask of abusive language detection. Therefore, to detect tweets containing hate speech, we explored all negative tweets found not to contain disinformation and filtered those that combined offensive/abusive terms as a preliminary stage before the identification of hate speech. This

[2] Term Frequency-Inverse Document Frequency (TF-IDF) is a statistical measure that indicates how many times a word/term appears in a piece of text.

differentiates our research from previous studies, which either relied on the personal judgment of the author to detect hate speech tweets (Assimakopoulos et al., 2017) or adopted a limited list of hate speech keywords (Aslan, 2017). Thus, we first identified offensive/abusive tweets based on extended prepared lists of Turkish words used in cyberbullying and offensive/abusive speech (Hüsünbeyi, Akar, & Özgür, 2022). Then, to spot the hate speech tweets, we explored the offensive/abusive tweets looking for specific entities that represent potential hate speech targets. These entities were derived from a collection of words/phrases that were either related to the religion/ethnicity of Syrians, targeting them based on gender identity, or inspired by the slang words/phrases (nicknames) used by the Turkish community to describe Syrians, such as "Suriyeli/Suriyeliler" (Syrian/Syrians), "Arap" (Arab), "Bedevi" (Bedouin), and "Çingene" (Gypsy).

Given that a tweet might include both disinformation and hate speech, we had the negative tweets subjected to a 2-phase classification process in which the disinformation tweets were first filtered and then further mined for hate speech content. The rest of the negative sentiment tweets, which we call "oppositional," as well as the "supportive" or positive tweets, were excluded from the analysis.

The models adopted for disinformation detection and hate speech recognition were evaluated based on a subset of 2,000 disinformation and 2,000 hate speech tweets (control group) manually annotated by two Turkish native speakers. Stratified sampling was used for disinformation and hate speech classification, similar to sentiment classification.

Findings and Discussion

The Main Discourse About Syrian Refugees on Turkish Twitter

Of the 33,100 tweets analyzed for sentiment, 89% (29,595) were negative, and 11% (3,505) were positive. The sentiment analysis model (Mulki et al., 2018) achieved a classification accuracy of 94%. The percentage of negative tweets against positive tweets reflects the findings of the existing literature, which points out the general negative sentiment against Syrian refugees on Turkish Twitter (Aridici, 2022; Ozduzen, 2020). It is also an illustration of the increasingly hostile environment against Syrians in Turkey (Farooq, 2021).

Although we used a limited number of trending hashtags as an initial basis for tweet collection (Table 1), the adopted 33,100 tweets contained 70,630 hashtags in total; of these, 510 hashtags were unique. We calculated the frequency of hashtags that were used more than 1,500 times. The top frequent hashtags are listed in Table 3.

Table 3. Most Frequent Hashtags in the Collected Tweets.

Hashtag (Turkish)	Hashtag (English)
#SuriyelilerinVatanıSuriyedir	*#Syrian's Home is Syria*
#KardeşimeDokunma	*#Don't Hurt My Brother*
#VatanEldenGidiyor	*#Our Homeland is Gone*
#afganlariistemiyoruz	*#We don't Want Afghans*
#AfganlarıAlmayın	*#Don't Host the Afghans*
#Suriyelileriİstemiyoruz	*#We don't want Syrians*
#ÜlkemdeMülteciİstemiyorum	*#I don't Want Refugees in My Country*
#SığınmacılarSınırDışıEdilecek	*#Asylum Seekers Will Be Deported*
#SuriyelilerSuriyeye	*#Syrians! Go to Syria*
#Müteciİstemiyoru	*#We don't Want Refugees*
#Altındağ	*#Altındağ (A Neighborhood in Ankara)*
#SuriyelileriAlmayın	*#Don't Host Syrians*
#YananHepBizOlduk	*#We're the ones who are burning*

Employing deictic words, such as "we," "my," and "our," the hashtags reproduced an exclusionary nationalist discourse, whereby Syrians were constructed as "other," as outsiders who do not belong in the national community. At the same time, Turkish people were constructed as victims, "the ones that are burning" and "whose homeland is gone." A few of the hashtags were explicitly associated with the August riots, such as #Altındağ, the area where hostilities against Syrians exploded, and #KardeşimeDokunma, used by promigrant groups during the riots.

We also explored how tweets of negative and positive opinions were distributed over time (see Figure 1).

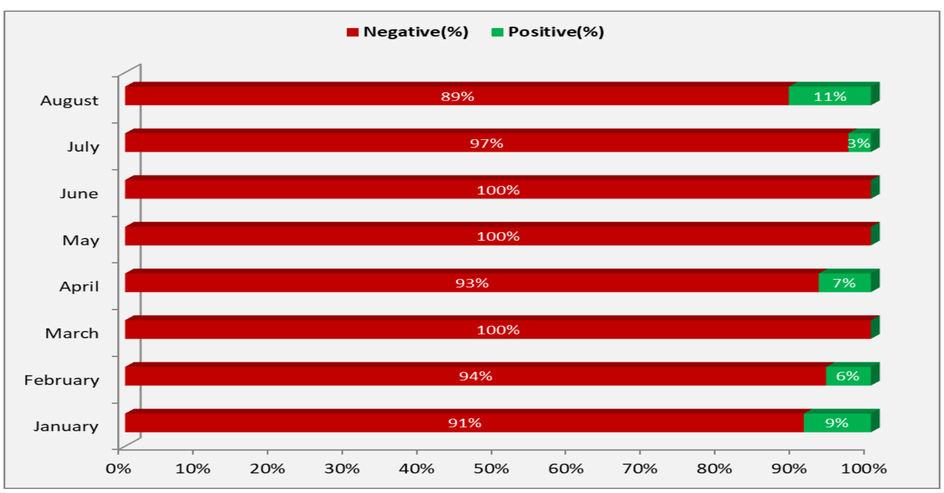

Figure 1. Monthly distribution (%) of positive/negative tweets.

Negative tweets dominated every month before the Muslim Eid holidays, when the Turkish government allowed Syrian refugees to visit their families in northern Syria. These short-term holidays fueled resentment among the Turkish community. Syrians returning to their homeland were used to arguing that Syria was safe for them. For example, Ümit Özdağ, Chairman of the Victory (Zafer) Party, addressed Syrians in a viral tweet:

> #ThisVisitHasLastedTooLong (#BuMisafirlikFazlaUzadı). This tweet is dedicated to the Syrian youth who went to Syria to celebrate Eid and get married. While they were going to the feast, the Turkish soldier who was either martyred in Syria and Mehmetçik or became disabled with his leg amputated, these soldiers were at the age of marriage. (Özdağ, 2021)

The biggest percentage of positive tweets, albeit still small at 11%, was spotted in August, the month of the violent episodes against Syrians in Ankara. It seems that support was only mobilized as a reaction to an extreme moment of violence and was only minimally expressed throughout the rest of the year. Given the politicization of the debate about Syrians, much of this support was expressed by government supporters.

Disinformation and Hate Speech Against Syrians

Having the 29,595 negative tweets subjected to specific NLP methods and lexical resources, as discussed in the methodology section, and with a classification accuracy of (91%), we identified 9,255 disinformation tweets, while the hate speech detection model recognized 9,374 hate speech tweets achieving an accuracy of (96%), while 1,077 tweets were found to have both disinformation and hate speech content. As the latter group of tweets was significantly small compared with each of the disinformation and hate speech tweet collections, it was excluded from the analysis. Similarly, we did not include "supportive" and "oppositional" tweets in further analysis. Table 4 reviews tweet examples of oppositional, disinformation, and hate speech categories, while Figure 2 illustrates the distribution of these tweet categories across the negative tweet collection.

Table 4. Examples of Oppositional, Disinformation, and Hate Speech Tweets.

Category	Tweet Example
Disinformation	Suriyeliler bombalandıkları için Türkiye'ye gelmiyorlar, Türkiye'ye gelmeleri için bombalanıyorlar. Emperyalizmin Orta Doğu'da oluşturmaya çalıştığı yeni yapay sınırlara Hayır demek için, #SuriyelilerSuriyeye! (personal communication, August 18, 2021) *Syrians do not come to Turkey because they are bombed, they are bombed to come to Turkey. To say "No" to the new artificial borders that imperialism is trying to create in the Middle East.* Benzine 55, motorine 67, LPG'ye 35 kuruş zam yapıldı. Türk halkı açlıktan çıldırıp intihar ederken, Suriyelilere 5 yıldızlı bahçeli ev yapılıyor. Türk halkı maruz kaldığı zamlarla Suriyelilerin yemesini, içmesini ve ev sahibi olmasını sağlıyor. Yeter artık! (personal communication, May 20, 2021) *Gasoline was raised by 55 cents, diesel by 67 cents, and LPG by 35 cents. While the Turkish people are going crazy from hunger and committing suicide, Syrians are being given 5-star houses with gardens. With the hikes Turkish people are subjected to, Syrians can eat, drink, and own a house. Enough is enough!*
Hate Speech	Araplar ayrı, bedeviler ayrıdır. Bizim ülkemizde arap sığınmacı yok, çöl bedevileri var. Bedeviler, dünyanın en pis ve barbar insanlarıdır Zamanında, küçük bir ücret karşılığında Peygamberi bile taşlamislardır. Esad,suriyeli bedevileri ülkemize pompalamıştır #KardeşimeDokunma (personal communication, August 12, 2021) *Arabs are one thing, Bedouins are another. There are no Arab refugees in our country, there are Bedouins from the desert. Bedouins are the most filthy and barbaric people in the world. In their time, they even stoned the Prophet for a small fee. Assad has pumped Syrian Bedouins into our country #DoNotTouchMyBrother* Ben ırkçı falan değilim! Nankör (arap) ve hain (arap) sevmiyorum! #SuriyelilerMutlu (personal communication, August 9, 2021) *I am not a racist! I don't like ungrateful (Arabs) and traitors (Arabs)! #SyriansAreHappy*
Oppositional	Ev içinde ev, devlet içinde devlet olmaz. Suriyelilerin milli varlığımız ve birliğimiz için tehdit oluşturmalarına izin vermeyeceğiz. Onun içindir ki 👇 #SuriyelilerSuriyeye #VatandaşlıkVermeYolVer (personal communication, May 12, 2021) *There cannot be a house within a house and a state within a state. We will not allow Syrians to pose a threat to our national existence and unity. For this reason, 👇 #SyriansToSyria #CitizenshipGrantingFacilitation*

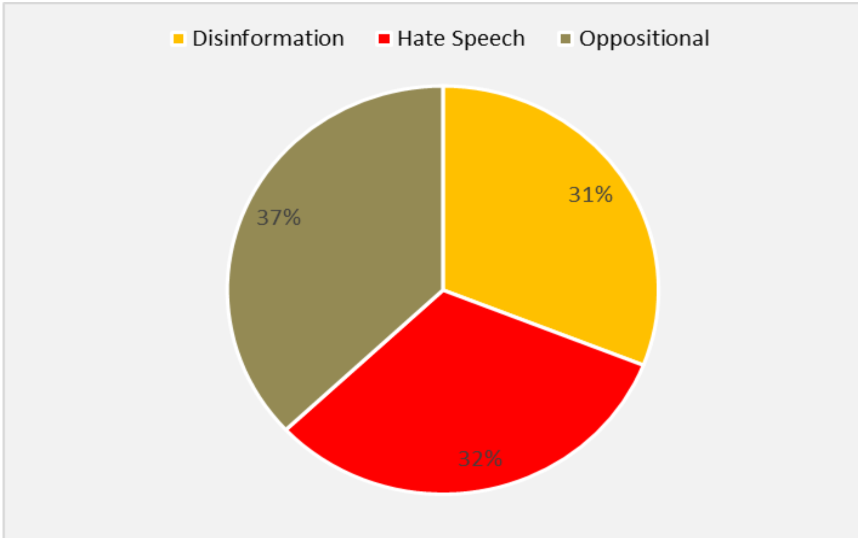

Figure 2. Distribution of disinformation, hate speech, and oppositional tweets.

As seen in Figure 2, the proportion of tweets containing disinformation and hate speech tweets within the negative tweet collection confirms the interrelation between negative sentiment toward Syrian refugees and the dissemination of disinformation and hate speech. Although platform restrictions about hateful conduct do not apply to the tweets expressing opposition to Syrian refugees as such—what we call here "oppositional"—the high volumes of hate speech and disinformation are disconcerting, given Twitter's efforts to battle both phenomena (Twitter, 2023) and the real-life impact these can have, inciting hate crimes and violence. Although this can be partially justified by the fact that Twitter algorithms cannot capture low-resourced languages, such as the Turkish language, for automatic toxic content removal, it is also revealing the lack of resources the platform is willing to allocate to this specific threat (Mulki, Haddad, Bechikh Ali, & Alshabani, 2019). For disinformation, the most common claims made in our data sample are listed in Table 5.

Table 5. Common Claims Extracted From Disinformation Tweets.

Top Frequent Claims
Suriyeliler Devleti kuruluyor
Syrians establish their country in Turkey
Mülteciler değiller
They are not refugees
Stratejik göç yapıyorlar
They perform a systematic migration
Su bedava
Water service is free (i.e., Syrians don't pay water bills)
Vergi yok
No taxes for Syrians

Some of these disinformation claims reproduced the myth of financial benefits enjoyed by Syrians (Aydınlı, 2020), constructing them as a financial burden to Turkey and comparing them to Turkish citizens who do not have access to such benefits. Others denied their status as refugees, claiming that their move to Turkey was deliberate and calculated rather than a desperate attempt to escape the war.

Similarly, when exploring tweets recognized as hate speech, we identified the most frequent words/terms listed in Table 6. Most of these words were used in the context of threatening and dehumanizing Syrians, mocking their culture, religion, and ethnicity as Arabs, while claiming the superiority of Turkish people over Syrians.

Table 6. Top Frequent Terms in Hate Speech Tweets.

Top Frequent Terms
Türküm (I am Turkish)
Arap (Arabs)
Türkiye Türklerindir (Turkey is for Turks)
Pislik (Dirt)
Irkçıyım (I am racist)
Bedevi (Bedouin)
Piçler, şerefsizler (Bastards)
Saldırısına (To attack)

About the distribution of disinformation and hate speech throughout the studied timeframe, as Figure 3 illustrates, throughout the studied period, the percentages of disinformation tweets have been consistently high, ranging between 49% and 80%. These tweets can be seen as part of disinformation campaigns against Syrians, often politically motivated by some opposition parties that adopt an antirefugee stance, such as the Zafer Party (Özerim & Tolay, 2021), as well as the broader disinformation that seems to shape the information ecosystem of an increasingly polarized Turkey (Karabat, 2018). On the other hand, hate speech tweets constituted nearly 40% of the tweets in January. A couple of Syrian-related news stories preoccupied the public agenda that month. First, the Ministry of Interior published the number of Syrians expected to voluntarily return to their country by the end of 2021. At the same time, an attack against a Syrian family took place in Izmir in mid-January. Although hate tweets did not exceed 20% between February and July 2021, they represented more than half (51%) of the negative tweets against Syrians during August, when the Ankara attacks against Syrians took place. This can be seen as an illustration of the close relationship between hate speech and violence, given that the former seems to have been instigated and likely further fueled the mobs against Syrian refugees in the August Altındağ riots.

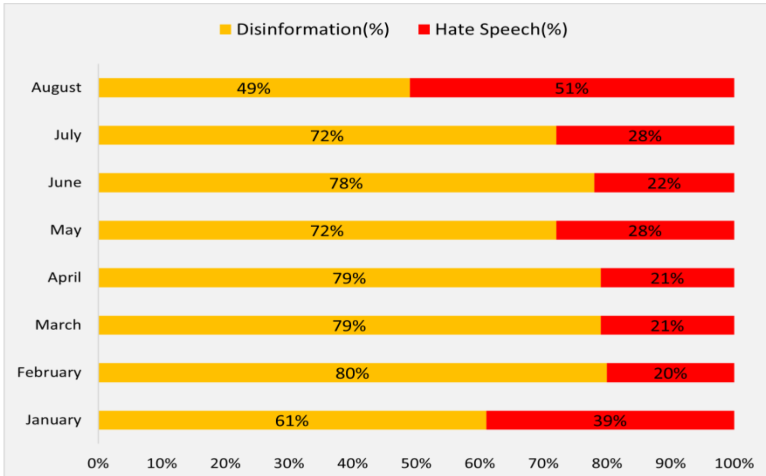

Figure 3. Monthly distribution of disinformation and hate speech tweets.

Who Tweets Against Syrians on Turkish Twitter?

Of the 12,177 tweets that included geolocation information in our sample, 10,701 tweets were associated with real locations. Figure 4 shows the distribution of the tweets across the international source locations, where we can see that the tweets posted from Turkey formed 90% of the investigated tweets. This is, of course, to be expected, given that many Turkish Twitter users reside in Turkey. The rest of the tweets were posted from countries in Europe, Canada, and the United States where Turkish and Syrian ex-pats resided.

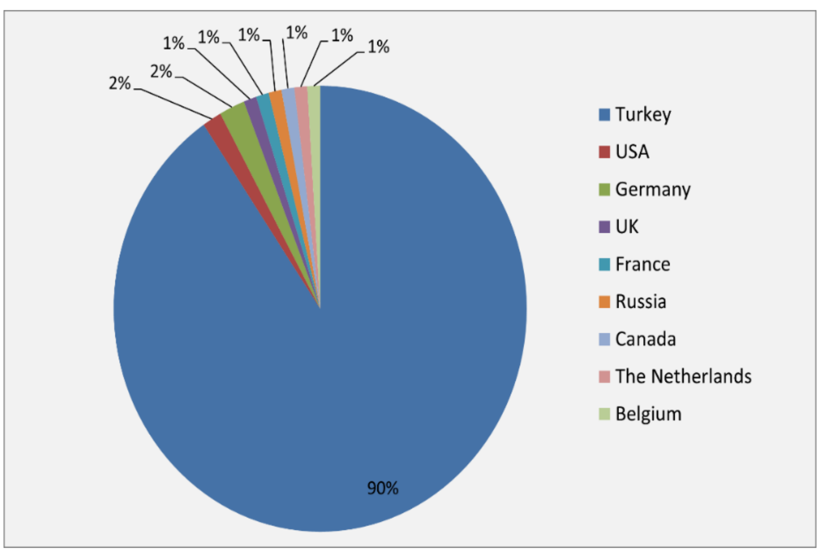

Figure 4. Tweet distribution among international locations.

Figure 5 illustrates how tweets in our sample were distributed among Turkish cities. In this chart, it is observed that Istanbul, Ankara, and Izmir were the most prominent in our sample as geopolitical locations. This is also to be expected, as these are the three major cities, as well as important decision-making centers. However, although most Syrian refugees resided in these three cities (UNHCR, 2021), an important portion of the tweets with Turkish locations (26%) were posted from different cities across the country. This indicates that the debates about Syrians within the Turkish community are not necessarily related to the distribution of Syrian refugees across Turkish cities but are an issue of general concern.

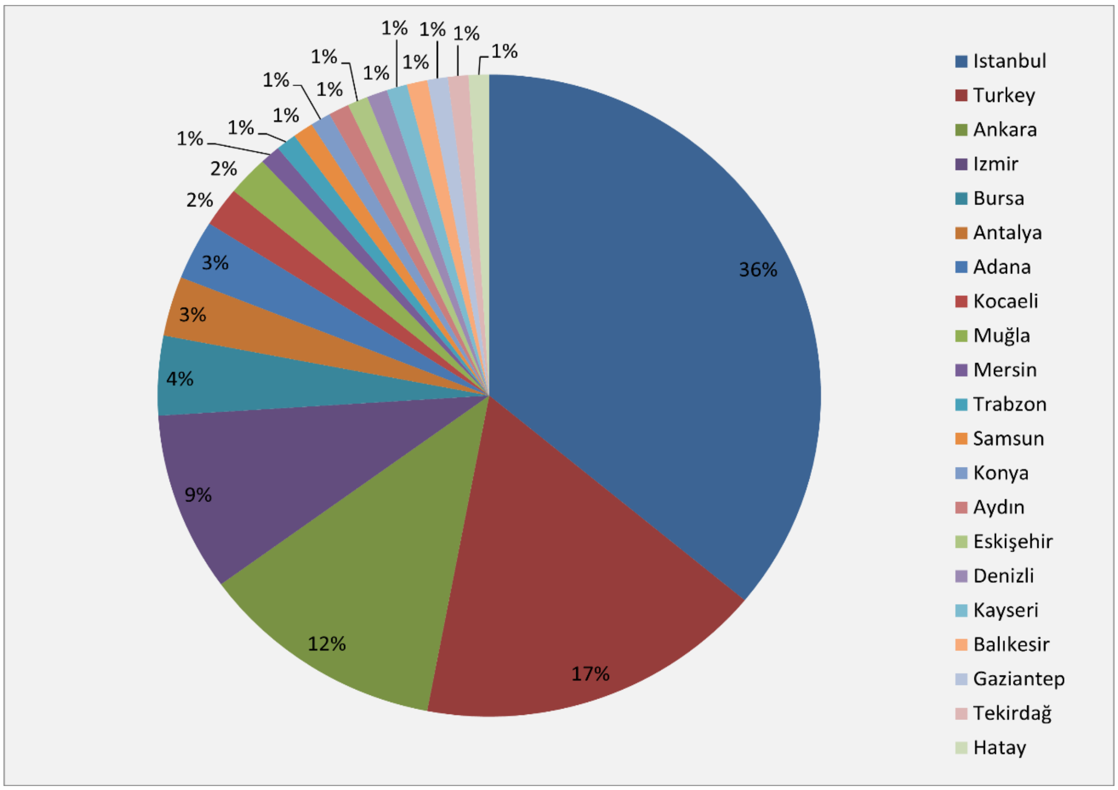

Figure 5. Tweet distribution among Turkish locations.

We further identified the source geolocations of disinformation and hate speech tweets. As Figure 6 shows, the distribution of disinformation and hate speech among international geolocations indicates that there are two countries where hate speech against Syrians seems to be higher than in Turkey, namely Germany and Russia, which can be explained on geopolitical grounds. Turkish people constitute Germany's largest minority at 3 million, importing a lot of national political tensions. At the same time, Russia's support for the Assad regime was expressed against Syrian refugees. Figure 7 shows the distribution of disinformation and hate speech tweets across the source geolocations in Turkey.

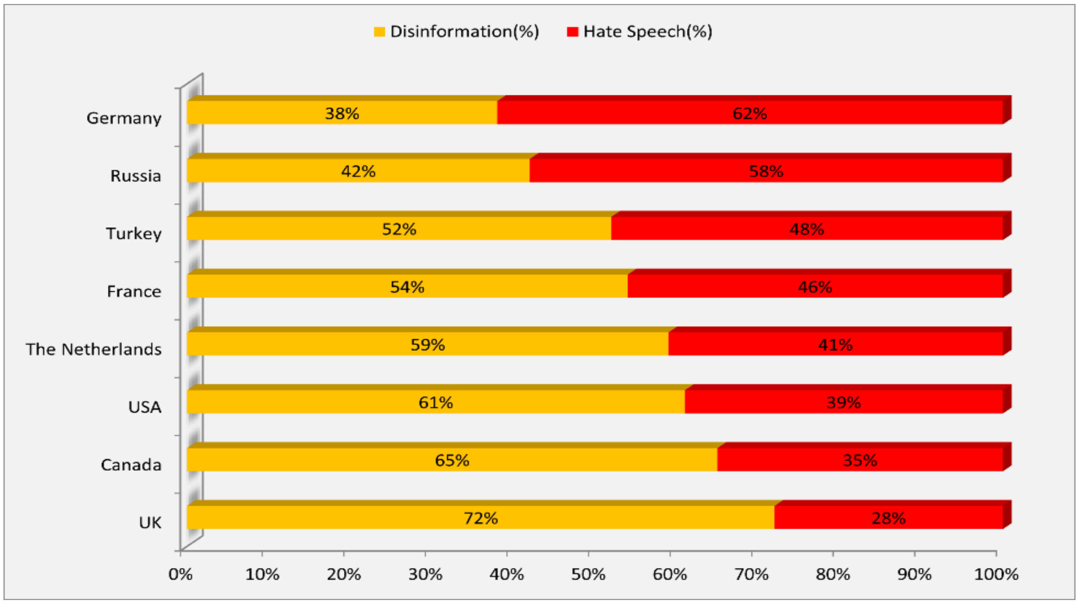

Figure 6. Disinformation/hate speech distribution for international locations.

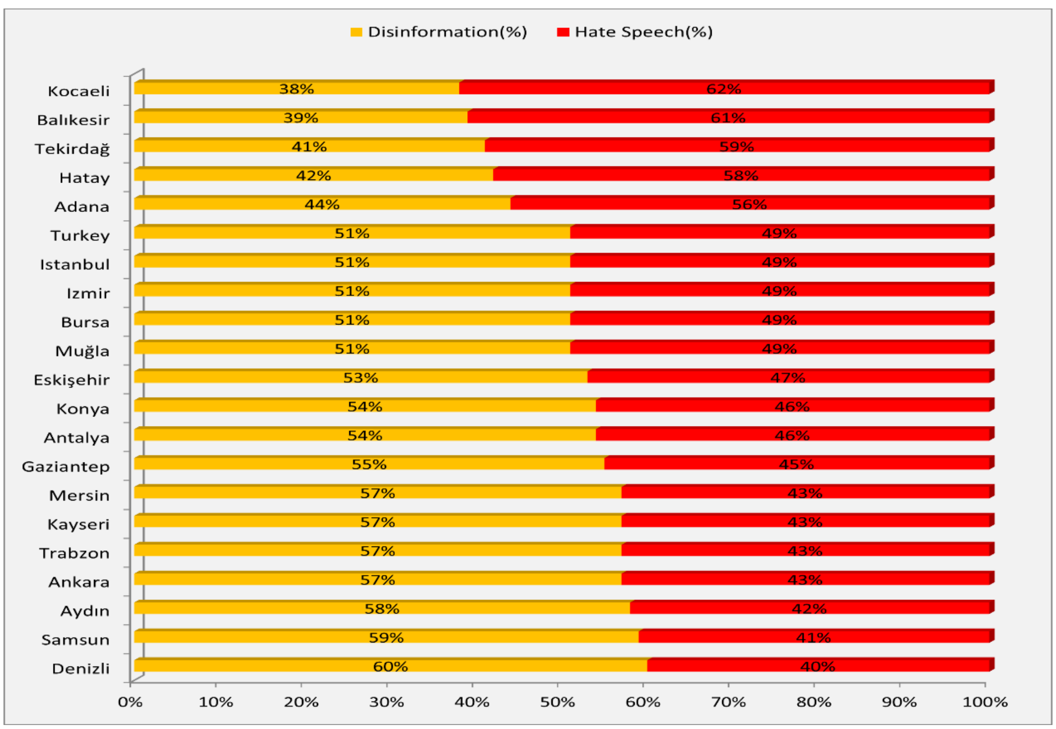

Figure 7. Disinformation/hate speech tweet distribution in Turkish locations.

It seems that hate speech was more prominent in the areas where Syrians were discussed the least, as seen in Figure 5.

Most importantly, given that the retweet rate achieved for all the tweets on a user's timeline is considered among the factors that identify how influential this user is, we tracked the tweeting and retweeting activity among the users in the disinformation and hate speech tweets to spot the influential users whose tweets were retweeted the most. The graphs below were produced by Gephi. The nodes denote the accounts involved in tweeting and retweeting, while the edges indicate that two users are related to each other by the retweet activity (i.e., one retweeted a tweet of the other). The colored clusters represent the communities of influencers and their retweeting users. These communities were created using a modularity-based clustering algorithm. In each community, we focused on the out-degree centrality that indicates the retweet rate of a user in this community, such that nodes of higher out-degree appear as bigger-sized nodes and denote the influential users, while the surrounding nodes connected to it represent the retweeting users. This was applied to disinformation and hate speech tweet collections to highlight the influencers and their communities of users who share the mentality/affiliation within which propaganda and polarized information spread well. Figure 8 shows the network of disinformation tweets with 6,385 accounts and 7,713 interactions.

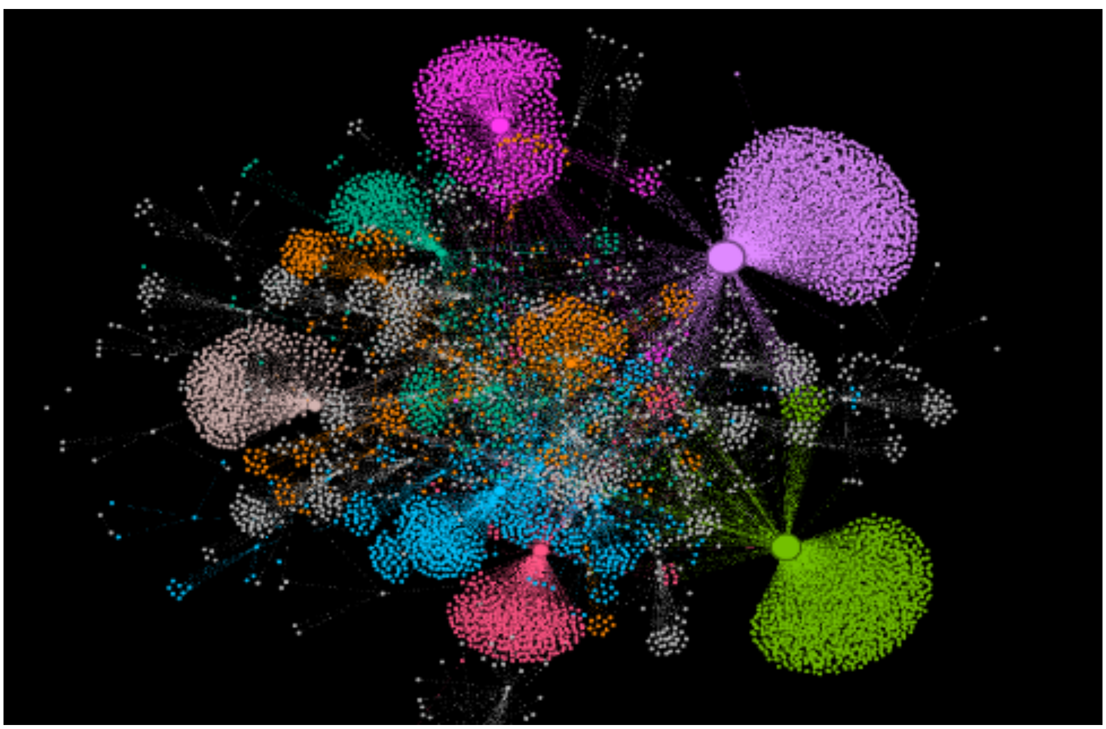

Figure 8. Disinformation tweeting/retweeting network.

What is evident from the network analysis is that disinformation circulates among a network of a few influential users, which are discerned here as bigger-sized nodes. As seen in Figure 8, influential users have a community of retweeters who amplify disinformation against Syrians by retweeting their tweets. What is notable, however, is that these influential users do not share followers or retweeters, which is because of their different political affiliations within the spectrum of oppositional parties. Indeed, when exploring the accounts of the influential users involved in spreading disinformation, as shown in Figure 9, we can see that the list of the top eight most-retweeted users included politicians, social media activists, and journalists either affiliated with the Victory (Zafer) Party, Good (İyi) Party and CHP or represented nongovernmental organizations and news agencies that adopt the Turkish nationalist line, such as the Yeni Çağ newspaper. These actors are representative of different parts of the political spectrum and are therefore followed and retweeted by different people.

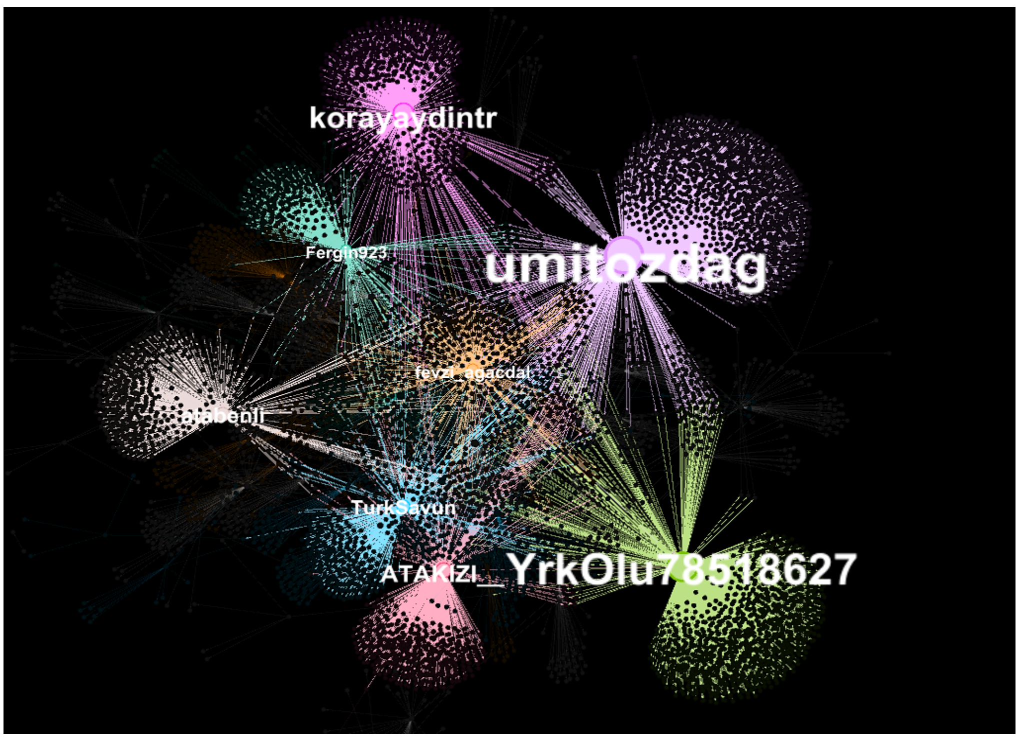

Figure 9. Most influential users in disinformation tweets.

At the same time, while 62% of the influential users had verified accounts and millions of followers, the rest of the accounts (38%) were not verified and had a moderate or small number of followers. Hence, when it comes to spreading disinformation about Syrian refugees, both the author's profile/affiliation and the content of the tweet, irrespective of the verification of the author's identity, play an instrumental role in how viral this tweet might become through likes and retweets.

On the other hand, compared with the disinformation network, the hate speech network in Figure 10 is denser, with overlapping nodes and edges. Influential users here have many retweeting users in common, indicating a coordinated tweeting and retweeting activity.

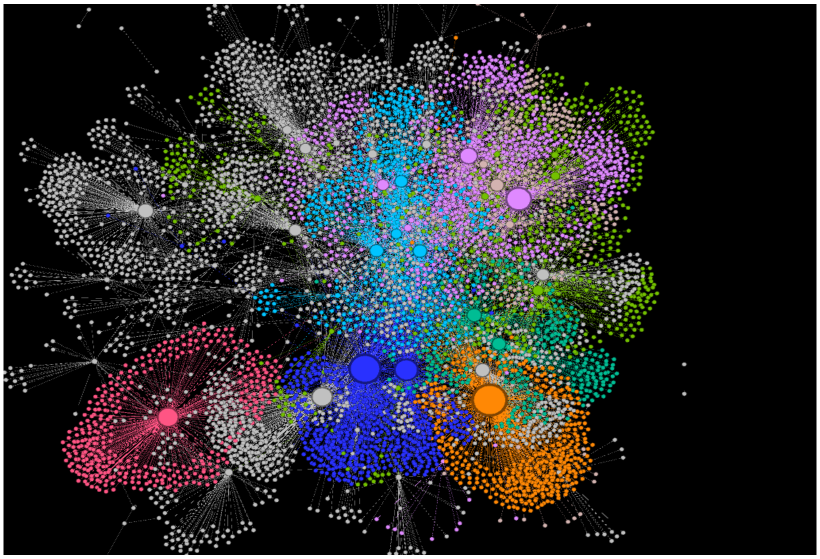

Figure 10. Hate speech tweeting/retweeting network.

Further supporting this conclusion is the fact that when exploring the accounts of the influential users after August 2021, we discovered that more than half the accounts involved in the hate speech network were either deactivated by their owners, removed by Twitter, or erased by the users (see Figure 11). If removal from Twitter can be seen as part of the platform's effort to tackle hate speech, the other two practices can be an indication of an orchestrated effort to instigate violence against Syrians. These tweets and related accounts seemingly had no reason to exist after the August 2021 attacks.

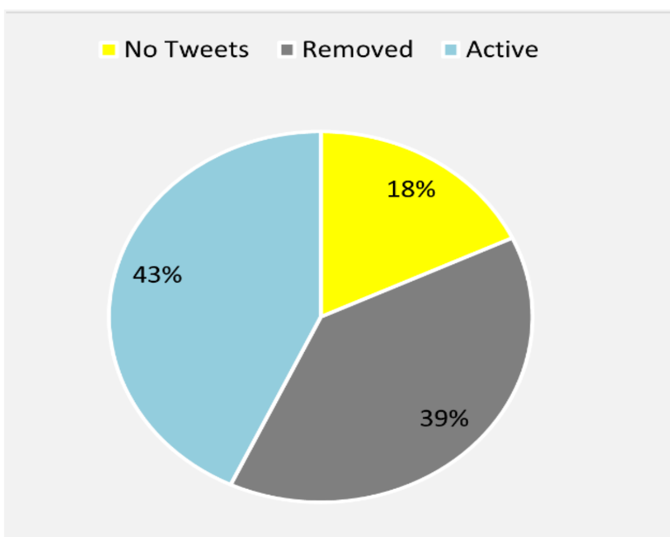

Figure 11. Activity status of hate speech influential accounts.

On the other hand, surprisingly, few influential users in the hate speech network shown in Figure 12 belong to politicians/former politicians affiliated with the opposition parties such as the Victory (Zafer) Party or CHP, while the rest of the influential party accounts seem to represent either ordinary Twitter users not affiliated with any political party or social media influencers (one of them was found to be supporting the AKP) besides journalists/activists who represent private media agencies and civil society associations such as Ülkü Ocakları (Grey Wolves) organization affiliated with MHP, People's Municipalities (Halkın Belediyeleri) news agency that echoes the ideas of both CHP and Good (İyi) Party, and The Silent Occupation (Sessiz İşgal) Youth Association and Voices from Home (Yurtten Sesleri) news agency, which is affiliated with Victory (Zafer) Party and CHP, respectively. The tweets/retweets made by most of these accounts have espoused the discourse of the opposition toward Syrians while spreading hate speech against them.

Moreover, when exploring their following/follower lists, we found that 66% of hate speech influential accounts were followers or friends of the disinformation-influential accounts identified in Figure 9, while 32% of hate speech influential users retweeted the tweets posted by disinformation-influential users.

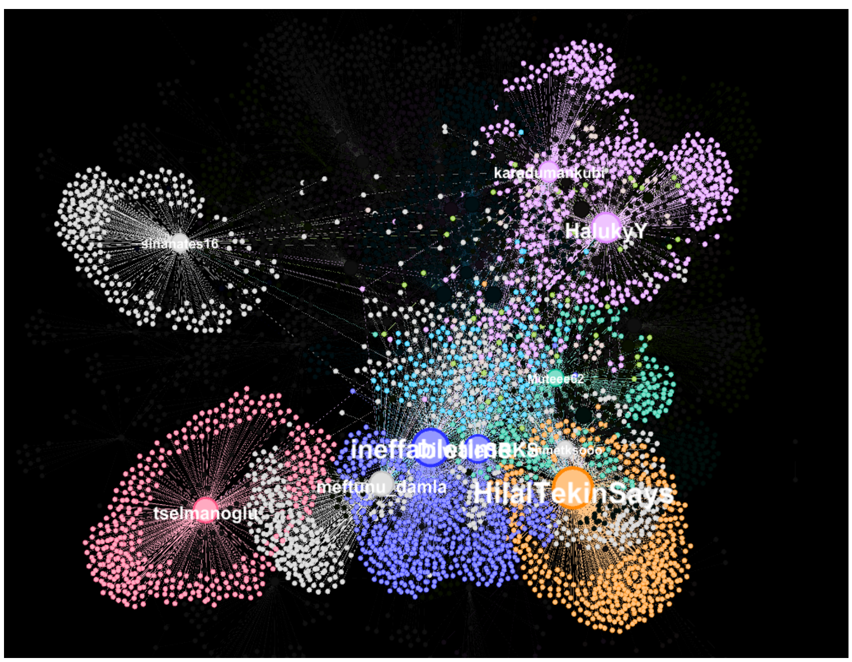

Figure 12. Most influential users in hate speech tweets.

The interaction between influential users in the hate speech network and disinformation network was further confirmed when merging the two networks, as shown in Figure 13, where the red clusters refer to the hate speech influential user communities, the yellow clusters indicate the disinformation-influential user communities and orange clusters/nodes denote the communities of users who were influential/retweeters in both disinformation and hate speech networks.

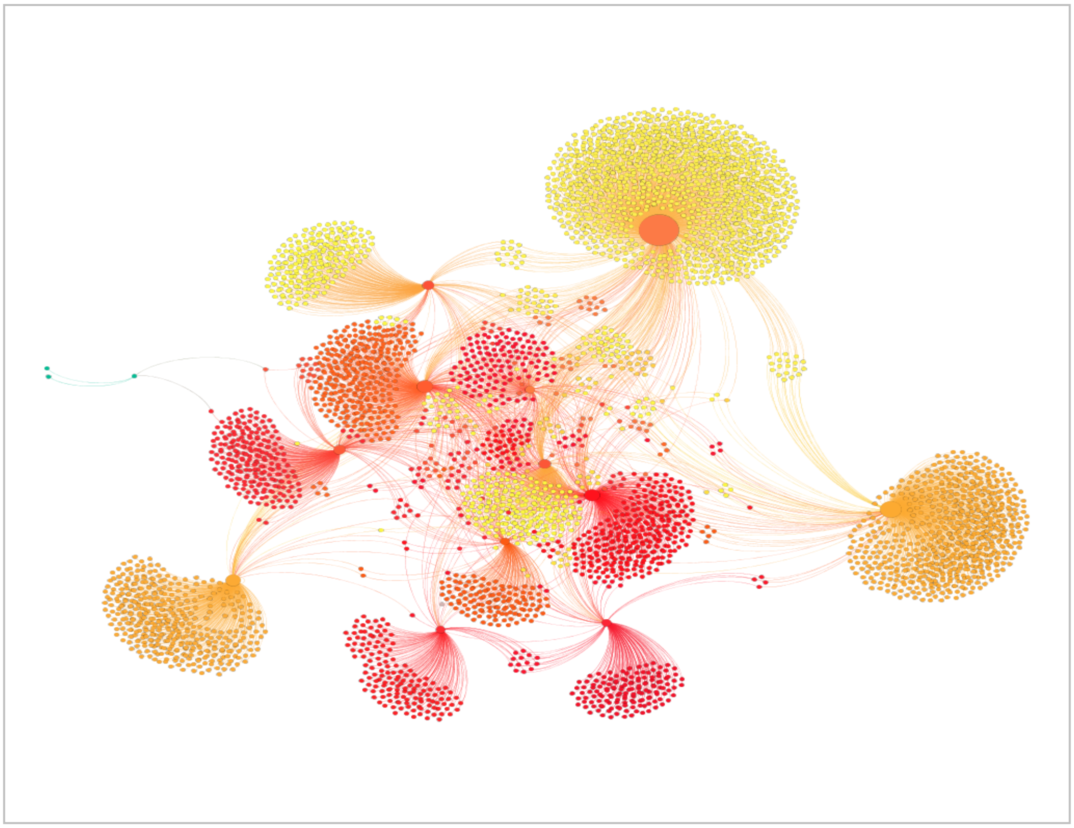

Figure 13. Disinformation-hate speech merged tweeting/retweeting network.

A closer look at the influential users within the merged hate speech and disinformation networks allows us to draw further conclusions. As seen in Figure 13, the influential users who were involved in hate speech were also influential users in the disinformation network and/or retweeters of disinformation-influential users. This confirms the interplay between disinformation and hate speech, illustrating that investigating hate speech actors and their propagation strategies cannot be conducted separately from the actors and amplification mechanisms identified in disinformation tweets. Although it is important to consider hate speech and disinformation as different discursive phenomena attacking Syrians, it is also crucial to see their interplay, which is crucial in reproducing online toxicity against refugees, ultimately inciting violent acts against them in consistency with specific political interests and agendas.

Finally, to further investigate the amplification of disinformation and hate speech within our sample, we performed a deep analysis of the accounts involved in retweeting the disinformation and hate speech content. This was done to check whether the propagation of disinformation and hate speech was spontaneous or further organized and propagated by fake accounts, the so-called bots. To this end, we first spotted the users whose retweets hit great values in each of the disinformation and hate speech tweets. Of the most retweeting users, we selected users whose account properties met those that describe bots, as bot accounts usually have few followers, their owners have recently joined Twitter, have no information in the bio section, or have no avatar or cover photo (Yang, Ferrara, & Menczer, 2022). Later, we fed the accounts identified from the previous steps into a machine-learning model trained to detect bot accounts.

As a result, we spotted 970 accounts as potential bots. All these users had recently joined Twitter, particularly right after the unrest between Turks and Syrians in Altındağ Ankara. Furthermore, most of these accounts (78%) had very few or no followers. Finally, all of these users followed the same accounts that were identified as influential users in either disinformation or hate speech tweets. Consequently, these accounts represent ideal bots. To confirm this, we subjected each of the potential bot accounts to an off-the-shelf application developed by Yang et al. (2022), where we found that out of the 970 studied accounts, 809 accounts (4% of the whole accounts) were identified as bots or fake accounts. Hence, it could be deduced that the anti-Syrian hashtags that were trending as part of disinformation and hate speech campaigns on Twitter were not just spontaneous expressions of public resentment. On the contrary, they were further supported by what seemed to be anti-Syrian propaganda, consisting of influential users and their followers from the bot account swarms. Figure 14 illustrates an example of a report produced for a bot account.

Figure 14. A sample result report for a bot account.

We found that 15% of the tweets were posted by bots. The impact of these fake accounts on disinformation, hate speech, and oppositional tweets is shown in Figure 15.

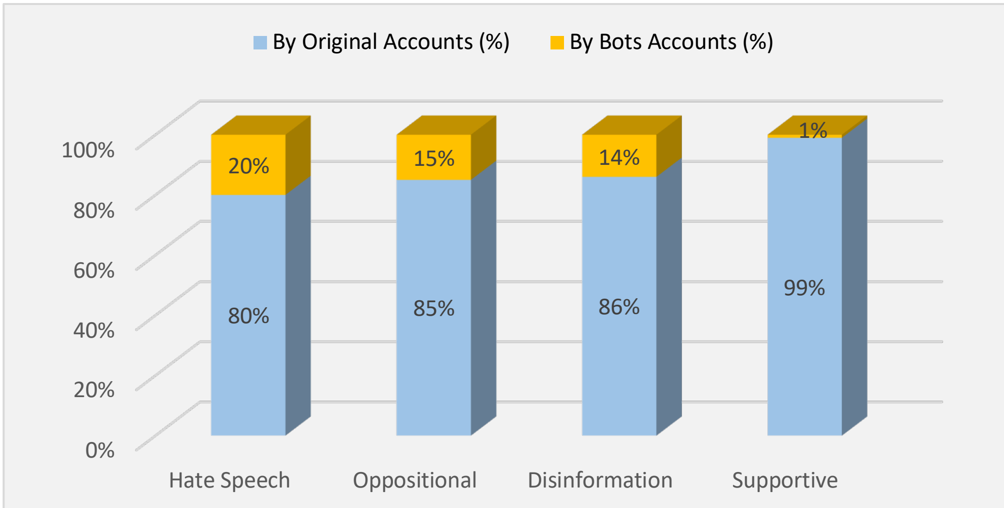

Figure 15. Percentages of tweets posted by bot accounts for each tweet category.

Conclusion

In this article, we investigated Syrian-related narratives on the Turkish Twittersphere. We went beyond classifying the attitudes toward Syrians to empirically show the distinction as well as the interplay between online hate speech and disinformation and how they are both encapsulated under the concept of online toxicity. This was practically conducted by introducing a novel framework that utilizes a machine learning–based sentiment analysis model along with NLP techniques to collect and investigate more than 30,000 Turkish tweets and recognize hate speech and disinformation tweets.

Our empirical findings construct a bleak image of the overall role of Twitter in the portrayal of Syrians in the Turkish Twittersphere. Despite the potential of social media as an alternative space of representation that can afford a voice to migrants and challenge hostile mainstream discourses, our research confirms earlier scholarship that illustrated social media as echoing rather than subverting negative portrayals of migrants and refugees. Employing the concept of online toxicity as a broader conceptual framework, we argue that it encompasses disinformation and hate speech as distinct but also interacting phenomena. We set out to study such negative discourses through a big data study of a sample of more than 30,000 tweets about Syrians in the Turkish language.

Our analysis proved the overwhelming presence of online toxicity against Syrians on Turkish Twitter. More than 30% of these negative tweets were examples of hate speech attacking Syrians and constructing them as inferior to the host population. A similar number of 31% among the negative tweets reproduced disinformation against Syrians, accusing them of unnecessarily leaving their homeland and ripping off the benefits provided in Turkey. The analysis indicated that there was an increase in hate tweets against Syrians during specific timeframes, namely, while the refugees were visiting Syria during the holidays and, most importantly, during the violent attacks against them in Ankara in August 2021. On the other hand, we believe that if we had enough tweets that contained both disinformation and hate speech, the interplay between hate speech and disinformation would have been investigated more deeply at the linguistic level in terms of how certain claims about Syrians could incite hatred against them.

At the same time, we also identified the Twitter networks circulating and amplifying these hate speech and disinformation tweets. We found that disinformation was most often instigated by political actors associated with the opposition, amplified by their followers, among which there was little overlap. Most of the hate speech actors in the network, on the other hand, either deleted their accounts, whether forcibly or voluntarily, or their relevant tweets attacking Syrians. There was considerable overlap between the remaining hate speech accounts and those (re)producing disinformation, indicating the interplay between the two forms of online toxicity, as spreading disinformation ultimately incited violent acts against Syrians, consistent with specific political interests and agendas. Finally, we identified that a considerable number (809) of accounts in our sample, responsible for (re)tweeting and spreading hate speech and disinformation, were bots. This, we argue, is another illustration of orchestrated attempts to undermine the presence of Syrians in Turkey.

Partially, these empirical findings are particular to Turkey, where the debate about Syrians has been highly politicized and has become a clear point of contention between the government and opposition parties. The intensity of these debates is such that politicians from the opposition are often the ones instigating disinformation against Syrians on Twitter. The background of this intense political polarization is a deepening economic crisis in Turkey, which has rendered Syrians evident scapegoats for populist politicians and the host population. The consequences of these populist arguments are experienced by Syrians, who are faced with increasing hostility in the country.

These findings also complement extant research in other countries that have illustrated social media as spaces that amplify and normalize negative sentiments against refugees and migrants (Aslan, 2017; Georgiou, 2018). However, the methodological design of our big data analysis allowed us to move beyond the critical analysis of media discourses and questions of representation. We illustrated how Turkish Twitter not only reflects and reproduces symbolic hierarchies of belonging that are ostensible in mainstream media and within Turkish society (Güney, 2022), but it does so in ways that are orchestrated and operationalized by influential social media actors. We, therefore, argue that to fully understand the dimensions and consequences of online toxicity and hate speech, we need to see them concerning disinformation and the ways in which they are politicized in specific sociopolitical contexts.

References

Akın, A. A., & Akın, M. D. (2007). Zemberek, an open source NLP framework for Turkic languages. *Structure, 10*(2007), 1–5.

Alkiviadou, N. (2019). Hate speech on social media networks: Towards a regulatory framework? *Information & Communications Technology Law, 28*(1), 19–35. doi:10.1080/13600834.2018.1494417

Alp, H. (2018). Suriyeli sığınmacılara yönelik ayrımcı ve ötekileştirici söylemin yerel medyada yeniden üretilmesi [The reproduction of discriminatory and othering discourse in local media directed toward Syrian refugees]. *Karadeniz Teknik Üniversitesi. Iletişim Fakültesi Elektronik Dergisi, 8*(1), 22–37.

Aridici, N. (2022). Constructing the 'national ideal': The 'inclusive' and 'exclusive' representations of Syrian refugees in Turkish print media. *Journal of Balkan and Near Eastern Studies, 24*(4), 696–714.

Aslan, A. (2017). Online hate discourse: A study on hatred speech directed against Syrian refugees on YouTube. *Journal of Media Critiques, 3*(12), 227–256. doi:10.17349/jmc117413

Assimakopoulos, S., Baider, F. H., & Millar, S. L. (2017). *Online hate speech in the European Union: A discourse-analytic perspective*. Cham, Switzerland: Springer International Publishing. doi:10.1007/978-3-319-72604-5

Aydınlı, S. (2020). *Hate and disinformation narratives in the online mediascape in Turkey* (No.12). South East European Network for Professionalization of Media (SEENPM). Retrieved from https://seenpm.org/wp-content/uploads/2020/12/Resilience-research-publication-2-Turkey-English.pdf

Az, İ., Gelişli, N., Barak, R., & Arslan, Z. (2017). *Media watch on hate speech report: January-April 2015* (No. 418). Hrant Dink Foundation. Retrieved from https://hrantdink.org/attachments/article/418/Media-Watch-On-Hate-Speech-January-April-2015.pdf

Bozdağ, Ç. (2020). Bottom-up nationalism and discrimination on social media: An analysis of the citizenship debate about refugees in Turkey. *European Journal of Cultural Studies, 23*(5), 712–730. doi:10.1177/1367549419869354

Chouliaraki, L., & Stolic, T. (2017). Rethinking media responsibility in the refugee 'crisis': A visual typology of European news. *Media, Culture & Society, 39*(8), 1162–1177. doi:10.1177/0163443717726163

Erdogan-Ozturk, Y., & Isik-Guler, H. (2020). Discourses of exclusion on Twitter in the Turkish context: #ülkemdesuriyeliistemiyorum (#idontwantsyriansinmycountry). *Discourse, Context & Media, 36*, 1–23. doi:10.1016/j.dcm.2020.100400

Farooq, U. (2021, November 3). Uncertainty for Syrians in Turkey as opposition warms to Assad. *Al Jazeera English*. Retrieved from https://www.aljazeera.com/news/2021/11/3/uncertainty-for-syrians-in-turkey-as-opposition-warms-up-to-assad

Georgiou, M. (2018). Does the subaltern speak? Migrant voices in digital Europe. *Popular Communication, 16*(1), 45–57. doi:10.1080/15405702.2017.1412440

Güney, Ü. (2022). Syrian refugees between Turkish nationalism and citizenship. *Journal of Immigrant & Refugee Studies, 20*(4), 519–532. doi:10.1080/15562948.2021.1950256

Hameleers, M., van der Meer, T., & Vliegenthart, R. (2021). Civilized truths, hateful lies? Incivility and hate speech in false information—Evidence from fact-checked statements in the US. *Information, Communication & Society, 25*(11), 1596–1613. doi:10.1080/1369118X.2021.1874038

Henry, S., Cuffy, C., & McInnes, B. T. (2018). Vector representations of multi-word terms for semantic relatedness. *Journal of Biomedical Informatics, 77*, 111–119. doi:10.1016/j.jbi.2017.12.006

Hübl, F., Cvetojevic, S., Hochmair, H., & Paulus, G. (2017). Analyzing refugee migration patterns using geo-tagged tweets. *International Journal of Geo-Information, 6*(10), 1–23. doi:10.3390/ijgi6100302

Hüsünbeyi, Z. M., Akar, D., & Özgür, A. (2022). Identifying hate speech using neural networks and discourse analysis techniques. In *Proceedings of the First Workshop on Language Technology and Resources for a Fair, inclusive, and Safe Society within the 13th Language Resources and Evaluation Conference* (pp. 32–41). Marseille, France: European Language Resources Association.

Idiz, S. (2019, January 10). Debate over Syrian refugees gathers steam in Turkey. *Al-Monitor*. Retrieved from https://www.al-monitor.com/originals/2019/01/turkey-syria-refugee-debate-gathers-steam.html

Karabat, A. (2018, December 14). *How Syrian refugees became victims on Turkey's "fake news" front line*. Retrieved from https://www.middleeasteye.net/news/how-syrian-refugees-became-victims-turkeys-fake-news-front-line

Karakas, B. (2021, August 18). *Refugees in Istanbul: Is anti-migrant sentiment growing in Turkey*. Deutsche Welle. Retrieved from https://www.dw.com/en/refugees-in-istanbul-is-anti-migrant-sentiment-growing-in-turkey/a-58890282

Kim, J. W., Guess, A., Nyhan, B., & Reifler, J. (2021). The distorting prism of social media: How self-selection and exposure to incivility fuel online comment toxicity. *Journal of Communication, 71*(6), 922–946. doi:10.1093/joc/jqab034

Kreis, R. (2017). #refugeesnotwelcome: Anti-refugee discourse on Twitter. *Discourse & Communication, 11*(5), 498–514. doi:10.1177/1750481317714121

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*. New York, NY: Data & Society Research Institute. Retrieved from https://datasociety.net/library/media-manipulation-and-disinfo-online/

Masullo Chen, G., Muddiman, A., Wilner, T., Pariser, E., & Stroud, N. J. (2019). We should not get rid of incivility online. *Social Media + Society, 5*(3), 1–5. doi:10.1177/2056305119862641

Mondal, S. K., Sahoo, J. P., Wang, J., Mondal, K., & Rahman, M. M. (2022). Fake news detection exploiting TF-IDF vectorization with ensemble learning models. In *Proceedings of Advances in Distributed Computing and Machine Learning* (pp. 261–270). Singapore: Springer Publishing Company. doi:10.1007/978-981-16-4807-6_25

Mulki, H., Ali, C. B., Haddad, H., & Babaoğlu, I. (2019). Tw-StAR at SemEval-2019 task 5: N-gram embeddings for hate speech detection in multilingual tweets. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 503–507). Minneapolis, MN: Association for Computational Linguistics. doi:10.18653/v1/S19-2090

Mulki, H., Haddad, H., Ali, C. B., & Alshabani, H. (2019). L-hsab: A Levantine Twitter dataset for hate speech and abusive language. In *Proceedings of the Third Workshop on Abusive Language Online* (pp. 111–118). Florence, Italy: Association for Computational Linguistics. doi:10.18653/v1/W19-3512

Mulki, H., Haddad, H., Ali, C. B., & Babaoğlu, İ. (2018). Preprocessing impact on Turkish sentiment analysis. In *Proceedings of the 2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1–4). Izmir, Turkey: IEEE. doi:10.1109/SIU.2018.8404718

Nockleby, J. T. (2000). Hate speech. *Encyclopedia of the American Constitution, 3*(2), 1277–1279.

Özdağ, Ü. [umitozdag]. (2021, July 26). *Suriye'ye hem bayramlaşmaya hem evlenmeye giden Suriyeli gençler için #BuMisafirlikFazlaUzadı Onlar bayramlaşmaya giderken, Suriye'de şehit olan Türk askeri de bacağı koparak gazi olan Mehmetçik de evlenecek yaştaydı....* [#ThisVisitHasLastedTooLong. This tweet is dedicated to the Syrian youth who went to Syria to celebrate Eid and get married. While they were going to the feast, the Turkish soldier who was either martyred in Syria and Mehmetçik or became disabled with his leg amputated, these soldiers were at the age of marriage…]. [Tweet]. Retrieved from https://twitter.com/umitozdag/status/1419756232873848835

Ozduzen, O. (2020). 'We are not Arabs and Taksim is ours' YouTubed political agency, place-making and social exclusion of Syrians. *City, 24*(5–6), 741–758. doi:10.1080/13604813.2020.1833538

Özerim, M. G., & Tolay, J. (2021). Discussing the populist features of anti-refugee discourses on social media: An anti-Syrian hashtag in Turkish Twitter. *Journal of Refugee Studies, 34*(1), 204–218. doi:10.1093/jrs/feaa022

Secen, S., Serhun, A., & Bekir, A. (2023). *Electoral dynamics, nationalist ideology, and party positions on Syrian refugees in Turkey*. Retrieved from https://doi.org/10.31235/osf.io/dbzfm

Şen, F. (2017). Bir 'Öteki' olarak mülteciler: Suriyeli mültecilerin anaakim ve alternatif medyada temsili [As 'the others' refugees: Representation of Syrian refugees in mainstream and alternative media]. *Atatürk İletişim Dergisi, 12*(1), 27–42.

Sheth, A., Shalin, V. L., & Kursuncu, U. (2021). Defining and detecting toxicity on social media: Context and knowledge are key. *Neurocomputing, 490*, 312–318. doi:10.1016/j.neucom.2021.11.095

Siapera, E. (2019). Refugee solidarity in Europe: Shifting the discourse. *European Journal of Cultural Studies, 22*(2), 245–266. doi:10.1177/1367549418823068

Stahl, K. (2018). Fake news detection in social media. *California State University Stanislaus, 6*, 4–15.

Statista Search Department. (2023, July 14). *Leading Twitter accounts in Turkey as of July 2023, by number of followers (in millions)* [Infographic]. Statista. Retrieved from https://www.statista.com/statistics/1318367/turkey-twitter-accounts-by-follower-numbers/

Turkish capital reels from violent protests against Syrians. (2021, August 12). *British Broadcasting Corporation News (BBC-News)*. Retrieved from https://www.bbc.com/news/world-europe-58185612

Twitter. (2023, April). *Synthetic and manipulated media policy*. Retrieved from https://help.twitter.com/en/rules-and-policies/manipulated-media

Udwan, G., Leurs, K., & Alencar, A. (2020). Digital resilience tactics of Syrian refugees in the Netherlands: Social media for social support, health, and identity. *Social Media+ Society, 6*(2), 1–11. doi:10.1177/2056305120915587

United Nations High Commissioner for Refugees (UNHCR). (2021). *UNHCR Turkey* [Fact sheet]. Retrieved from https://data.unhcr.org/en/documents/details/88665

Vidgen, B., & Yasseri, T. (2020). Detecting weak and strong Islamophobic hate speech on social media. *Journal of Information Technology & Politics, 17*(1), 66–78. doi:10.1080/19331681.2019.1702607

Wang, S., Dai, Y., Shen, J., & Xuan, J. (2021). Research on expansion and classification of imbalanced data based on SMOTE algorithm. *Scientific Reports, 11*(1), 1–11. doi:10.1038/s41598-021-03430-5

Waseem, Z., Davidson, T., Warmsley, D., & Weber, I. (2017). Understanding abuse: A typology of abusive language detection subtasks. In *Proceedings of the First Workshop on Abusive Language Online* (pp. 78–84). Vancouver, Canada: Association for Computational Linguistics. doi:10.18653/v1/W17-3012

Yang, K.-C., Ferrara, E., & Menczer, F. (2022). Botometer 101: Social bot practicum for computational social scientists. *Journal of Computational Social Science, 5*(2), 1511–1522. doi:10.1007/s42001-022-00177-5

Yücel, A. (2021). Symbolic annihilation of Syrian refugees by Turkish news media during the COVID-19 pandemic. *International Journal for Equity in Health, 20*(1), 1–11. doi:10.1186/s12939-021-01472