The Syrian War on Facebook: A call for mixed methods
by Toni Rouhana

Abstract
In this paper, I analyze 32 public Facebook pages that actively support a side in the ongoing Syrian turmoil to study users’ online expressions and test whether the war has translated metaphorically onto these Facebook pages. This study is in conversation with and contributes to the online political polarization scholarship, the role of online platforms in the “Arab Spring,” and the digital methods used to analyze online data in Arabic. I use a supervised, machine-learning approach to classify more than four million unique comments and more than 70,000 unique posts, from Facebook to check whether the Facebook pages are becoming warring spaces where opposing parties “fight.” Or whether they are producing spaces for conversations and “connections”? Or whether they are just echo chambers? My findings show that these pages are producing echo chambers rather than becoming warring spaces or conversing spaces. My findings also show that religious expressions are overwhelmingly used by all sides. I conclude by suggesting future methodological and analytical approaches for research of social network sites (SNS).

Contents
Introduction
Summary of the unfolding struggle in Syria
Related works
Argument for analysis of content rather than users
Methodology
Data classification
Results
Discussion
Limitations and future work
Conclusion

Introduction
In the wake of the “Arab Spring,” researchers took an interest in the role that social network sites (SNS) and online technologies played in these revolutions. In the immediate aftermath of the uprisings, scholarly
research, popular media, and even participants looked at social media as the driver and enabler of these revolutions (Alhindi, *et al.*, 2012; AlSayyad and Guvenc, 2015; Clarke and Koçak, 2019; Eltantawy and Wiest, 2011; Ghonim, 2012; Khamis, *et al.*, 2012). Debates about the role of the Internet in shaping social and political lives are certainly not new (Castells, 1996), but much scholarly work on this subject has flourished since the appearance and immediate popularity of SNS. Most importantly, since 2010, research has focused on the roles that SNS played in the organization, mobilization, and outreach of social movements [1], with a special interest in the “Arab Spring” and protests’ diffusion across nation-states with long-standing authoritarian regimes throughout the Arab world.

Two general questions animate this research: What aspects of daily life war related dynamics and interactions for ordinary Syrians have translated online through these platforms and what has not? And for those aspects that have been translated, how has that taken place, what is gained, what is lost?

This paper tests whether the war in Syria has translated metaphorically to Facebook and whether people are using public pages as virtual battlegrounds. I studied users’ comments on 32 Facebook public pages — that either exclusively cover the war or support one of the warring sides on the grounds or doing both — to test whether users used these spaces to virtually attack each other.

---

**Summary of the unfolding struggle in Syria**

Following Tunisia, Egypt, and Libya in what became known as the “Arab Spring,” Syria was one of the states that witnessed massive demonstrations that swept the country. At the beginning of the uprisings, Syrians went to the streets to demand economic and political reforms, the lifting of the emergency law in place since 1963, respect for basic human rights, and the reformation of the many police and security services branches that had surveilled, oppressed, and persecuted Syrians, especially those engaged in political opposition, for decades [2]. These demonstrations were met with brutality which, rather than crushing the uprisings and silencing the demonstrations, led to, on the one hand, the expansion of the demonstrations, and on the other hand, the emergence of armed resistance in many regions. By the end of 2011 what had started as peaceful demonstrations turned into a full-fledged civil war with sect-based divisions (Wimmen, 2016; Ogunnowo, *et al.*, 2020). The 10 years of war had devastating effects. At the time of this writing, the death toll estimate is between 350,000 and 500,000 (*New York Times*, Human Rights Watch). The United Nations stopped counting in 2016, due to the impossibility of verifying and documenting deaths in the ongoing war conditions. According to the United Nations High Commissioner for Refugees (UNHCR), the war also caused 5.6 million refugees to flee to neighboring countries and 6.6 million were internally displaced. The U.N. Economic and Social Commission for Western Asia (ESCWA) estimated the economic damage of the war to US$388 billion in 2018.

---

**Related works**

**Political polarization**

Studies of online political polarization focus primarily on user behavior and activities on either single or multiple online SNSs, mainly Facebook, Twitter, and YouTube. These studies implement multiple methodological approaches. Some use analysis of large-scale datasets of user connections and their online behavior (Howard, *et al.*, 2011; Bruns, *et al.*, 2013; Jacobson, *et al.*, 2016; Batorski and Grzywińska, 2018). Others use small-scale content analysis of a sample of users, or select pages from Facebook and Twitter (Eltantawy and Weist, 2011; O’Callaghan, *et al.*, 2014; Ernst, *et al.*, 2017). Yet others (Burkell, *et al.*, 2014) employ traditional qualitative methods that include surveys, interviews, and focus groups in order to
analyze users’ behavior and perception of social media. Still others (Theocharis and Lowe, 2016; Bail, et al., 2018) conduct experiments in order to study changes in users’ political alignment, views, and levels of activism.

All these studies culminate in user-centric results and a kind of social network analysis in order to classify users’ political/online affiliations. In other words, users’ accounts are classified as part of a group of users sharing similar political views [3]. Additionally, this kind of classification is dependent on users’ affiliations. For example, on Facebook, this classification often relies on the analysis of the pages that people follow, which dismisses comments that Facebook users leave on pages that they do not follow. Instead, this paper analyzes political polarization by using online content produced by users rather than analyzing the users’ networks. This approach opens up the possibility to go beyond the testing of political polarization and research the discursive processes that take place on SNSs.

Social media use during the Arab Spring

The role of social media in social movements, political polarization, and organizing is no longer questioned (Howard, et al., 2011; Salvatore, 2011) [4]. In the case of the Arab Spring in general and Syria in particular, most studies — whether large scale or not — use a social network classification of online accounts to draw connections between different users. Howard, et al. (2011) described three major roles played by social media in the Tunisian and Egyptian uprisings. First, they found that social media played a critical role in shaping political debates in the Arab Spring. Second, they found a massive increase in what they call “online revolutionary conversations” preceding events taking place off-line [5]. Last, they found that social media diffused democratic ideas across national borders. O’Callaghan, et al. (2014) proposed a new approach to studying the role of two major social media platforms (Twitter and YouTube) concerning the Syrian case study. They propose an exploratory, small-scale [6] study, free of the predetermined assumptions and hypotheses that usually characterize the study of online political polarization, in order to capture the complexity of the Syrian situation. In addition, their study aims to move beyond the pro-regime/opposition classification which they argued obscures the multiplicity of factions within the opposition. They classify users into four major clusters: Pro-Assad, Kurdish, Secular/Moderate, Jihadist. They also used a subgraph that represents the differences within the anti-Assad cluster, which shows three different communities that could be considered opposed to each other, revealing the nuances of the category of “opposition to Syrian regime.”

Lynch, et al. (2014) dubbed the Syrian conflict the most socially mediated in history [7]. Using a Twitter dataset that spans over 28 months, beginning on 1 January 2011, and ending on 30 April 2013, they first compared discourses in tweets written in English and in Arabic, arguing that English-language tweets focused on different topics than Arabic tweets [8]. They then explored the changes in influence between English and Arabic tweets based on the number of retweets. They demonstrated that English tweets dominated the top 250 tweets during 2011 while Arabic tweets concerned with Syria gained prominence later on. Lastly, they showed that Twitter accounts became clustered in more insular groups with time: “The cluster analysis demonstrated the shift from a fairly decentralized Syrian Twittersphere tightly embedded in the broader Arab Spring narrative into the consolidation of multiple, increasingly insular competing networks” [9]. This finding overlaps with previous research revealing that SNSs bring like-minded users together and produce “filter bubbles” and “echo chambers.”

The Syrian Civil War and studies of SNSs

In the case of Syria’s uprising and civil war, few studies have focused on the role of Facebook in the preparation, organization, and mobilization of the uprisings, with the exception of Al-Mustafá (2012). Al-Mustafá conducted a content analysis of three main Facebook pages which were directly linked to the Syrian revolution and that supported Syrian activists in 2011 and early 2012. Al-Mustafá argued that the Facebook pages took over the role of traditional mainstream media as a source of information in the public sphere and claimed that these pages shaped the revolutionary political discourse. In some cases, Al-Mustafá argued that some of the conversations that took place on these pages led to downturns and splits that took
The Syrian War on Facebook: A call for mixed methods

This book provides an invaluable study of the different approaches each of these Facebook pages implemented, in order to show the extent to which they influenced public opinion, helped organize the revolution, and mobilized people as well as created divisions within the opposition itself. Al-Mustafà, analyzed pages, posts, and comments throughout 2011 and showed how the narrative shifted from posts focused on freedom, democracy, and reform to a narrative centered on sectarian divisions and hate when the uprising started to develop into a civil war towards the end of 2011.

Other ways researchers have studied Facebook in the context of the Arab Spring, provide primarily a general analysis of the role of Facebook and are based on the researchers’ observations in the context of either a comparison with other “Arab Spring” revolutions (Khamis, et al., 2012), or as part of analyzing the opposing visual propaganda tactics between regime and opposition supporters (Seo and Ebrahim, 2016). Shehabat (2012) studied the media cyber-war occurring on different online platforms between the Syrian Electronic Army (SEA) and the Syrian Free Army and included a section on the tactics that the SEA used in order to shut down opposition Facebook pages and sometimes hack users’ accounts.

Argument for analysis of content rather than user

Analyzing online conversations rather than users’ connections produces data points of conditions which cause fluctuations in these divisions and highlights the magnitudes of local, national, and international events unfolding throughout the war in influencing sectarian divides. So rather than searching for sectarian people, I searched for conditions that intensified sectarian divisions based on online discourse. For example, following Rogers’ Digital Methods Initiative approach to study the Web as a cultural and social source of information, Ben David and Fernández (2016), combined a multimodal content analysis of text, images, and links to a dataset of Spanish right-wing political parties’ Facebook pages.

Using Mouffe and Laclau’s concept of agonism which opens a space for the possibility of democratic competition based on constructive adversarial competition rather than hostility, Sack (2004) visualized the possibility of establishing a dialogue with as many players as possible, regardless of the potential confrontational arguments that include a multiplicity of voices. His visualization shows that argument does not have to be war in response to George Lakoff and Mark Johnson’s position that argument is metaphorically war and therefore “a culture in which arguments are not conceptualized as verbal warfare, but as collaborative dances.”

In this study, I test whether the war translated into arguments on Facebook or whether this translation produced dialogue in the sense of agonistics which is what my ethnographic field work indicated was taking place, even if on a small scale, between different individuals and groups in Syria and Lebanon. Eighty out of 81 interviewees I interviewed between 2016–2019 said they believe that after the war, Syrians will live together regardless of the outcome of the war. They also affirmed that they are still in contact with friends and families that have opposing political views. They do not deny heated debates about the future of Syria, but they agreed that they could find a way to live together across political, social, and sectarian differences. For example, Ahmed, who actively protested since the early days of the Revolution and was tortured and jailed by the Syrian security forces and still has apparent torture marks on his body, is part of a collective that includes supporters of almost all political and military players in Syria. Like almost everyone I interviewed, Ahmed affirms that the collective was looking for a common ground to build a different future for themselves in Syria.

I focus on the conversations that took place on 32 politically engaged Facebook pages in Syria in the form of comments on posts published by these pages, rather than using a user-centered approach. In order to examine the dynamics of the conversations taking place in these venues, I analyzed the textual content of all Facebook posts and their respective comments associated with these Facebook pages (see Table 1). Building on my classification method (described later), I identified textual markers that could be used as
identifiers of each faction and of different sectarian communities online. I explored whether the online discourse contributed by users of these pages could be an indicator of a person’s political affiliation.

<table>
<thead>
<tr>
<th>Facebook Page ID</th>
<th>Facebook Page Name</th>
<th>Political Support</th>
<th>Number of Posts</th>
<th>Number of Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>11195192150190</td>
<td>NYCCCBB</td>
<td>Pro-regime</td>
<td>2,461</td>
<td>65,624</td>
</tr>
<tr>
<td>1416705415237220</td>
<td>syratwutreagency</td>
<td>Pro-regime</td>
<td>1,750</td>
<td>17,182</td>
</tr>
<tr>
<td>29371290802507</td>
<td>ALMESO83</td>
<td>Pro-regime</td>
<td>2,455</td>
<td>85,384</td>
</tr>
<tr>
<td>306800620173703</td>
<td>mouhibi83</td>
<td>Pro-regime</td>
<td>2,349</td>
<td>157,351</td>
</tr>
<tr>
<td>396319977108468</td>
<td>SanaNews</td>
<td>Pro-regime</td>
<td>998</td>
<td>6,123</td>
</tr>
<tr>
<td>461256323946384</td>
<td>Syria1VChannels</td>
<td>Pro-regime</td>
<td>2,320</td>
<td>76,132</td>
</tr>
<tr>
<td>481691578513721</td>
<td>samatv.net</td>
<td>Pro-regime</td>
<td>2,989</td>
<td>135,755</td>
</tr>
<tr>
<td>757847354293105</td>
<td>frsan.Alassad</td>
<td>Pro-regime</td>
<td>1,003</td>
<td>5,649</td>
</tr>
<tr>
<td>782503618434620</td>
<td>SYRIAN.SyriaRealInfosAndNews</td>
<td>Pro-regime</td>
<td>1,403</td>
<td>7,314</td>
</tr>
<tr>
<td>788757481176531</td>
<td>SMT.Official.Page</td>
<td>Pro-regime</td>
<td>1,258</td>
<td>11,035</td>
</tr>
<tr>
<td>107868985943631</td>
<td>AlDouma.tv</td>
<td>Pro-regime</td>
<td>2,831</td>
<td>113,142</td>
</tr>
<tr>
<td>332052706876392</td>
<td>ajnnewssyria</td>
<td>Pro-regime</td>
<td>993</td>
<td>3,434</td>
</tr>
<tr>
<td>208881799134874</td>
<td>aikhbaria.Sy</td>
<td>Pro-regime</td>
<td>3,580</td>
<td>121,562</td>
</tr>
<tr>
<td>1872679796022</td>
<td>AsmaaAlAssad.FirstLady</td>
<td>Pro-regime</td>
<td>2,518</td>
<td>213,547</td>
</tr>
<tr>
<td>14422664897809</td>
<td>DNNAR</td>
<td>Pro-regime</td>
<td>4,447</td>
<td>125,569</td>
</tr>
<tr>
<td>18408466126254</td>
<td>im syrian</td>
<td>Pro-regime</td>
<td>1,116</td>
<td>11,464</td>
</tr>
<tr>
<td>201464629874643</td>
<td>fatakia.news.network</td>
<td>Pro-regime</td>
<td>3,577</td>
<td>106,792</td>
</tr>
<tr>
<td>20468139550123</td>
<td>e.maher.assal.assal.loverson</td>
<td>Pro-regime</td>
<td>3,853</td>
<td>169,420</td>
</tr>
<tr>
<td>29188740416834</td>
<td>vescy</td>
<td>Neutral</td>
<td>405</td>
<td>1,010</td>
</tr>
<tr>
<td>469605269833613</td>
<td>YomytKazefeh</td>
<td>Neutral</td>
<td>3,722</td>
<td>1,296,937</td>
</tr>
<tr>
<td>111632495584341</td>
<td>Douma.Revolution.2011</td>
<td>Pro-Opposition</td>
<td>3,865</td>
<td>52,193</td>
</tr>
<tr>
<td>150398036591540</td>
<td>Ghazaleh.Center</td>
<td>Pro-Opposition</td>
<td>519</td>
<td>15,237</td>
</tr>
<tr>
<td>16578006805908</td>
<td>ShaaNetworkArabic</td>
<td>Pro-Opposition</td>
<td>4,484</td>
<td>469,433</td>
</tr>
<tr>
<td>17573795813435</td>
<td>hamza.alshafeed</td>
<td>Pro-Opposition</td>
<td>4,523</td>
<td>191,131</td>
</tr>
<tr>
<td>191741018515060</td>
<td>HadiAlaballah</td>
<td>Pro-Opposition</td>
<td>1,978</td>
<td>785,475</td>
</tr>
<tr>
<td>256265171116458</td>
<td>syria.breaking</td>
<td>Pro-Opposition</td>
<td>3,844</td>
<td>336,385</td>
</tr>
<tr>
<td>256307197580863</td>
<td>snhr</td>
<td>Pro-Opposition</td>
<td>1,006</td>
<td>3,113</td>
</tr>
<tr>
<td>426796315726</td>
<td>Syrian.Revolution</td>
<td>Pro-Opulsion</td>
<td>3,813</td>
<td>454,844</td>
</tr>
<tr>
<td>42091170794376</td>
<td>LensYoungHomsi</td>
<td>Pro-Opulsion</td>
<td>1,662</td>
<td>23,306</td>
</tr>
<tr>
<td>680758535289951</td>
<td>Ri'SmiediaOffice1</td>
<td>Pro-Opulsion</td>
<td>566</td>
<td>1,911</td>
</tr>
<tr>
<td>343172915751705</td>
<td>LensYoungdimashqi</td>
<td>Pro-Opulsion</td>
<td>3,678</td>
<td>45,051</td>
</tr>
<tr>
<td>326766683114</td>
<td>syriahro</td>
<td>Pro-Opposition</td>
<td>2,540</td>
<td>65,498</td>
</tr>
</tbody>
</table>

**Table 1:** Facebook pages from which data was collected.
Note: I have included Facebook Page IDs as a finding tool as some of the pages’ titles have changed or no longer exist.

Based on the overt affiliation of these pages two possibilities could be deduced: either these pages were becoming “battlegrounds” for opposing users to “attack” and supporters to “defend”, or they became echo chambers — if a page supported the regime, it created a space for regime supporters to interact and converse with each other rather than being an open space of conversation that included views from supporting, opposing, and neutral voices. This also applied to opposition pages. Last, I presented an initial comparison between findings from online data and from my eighteen-month ethnographic fieldwork to argue that online behavior did not necessarily reflect off-line dynamics, calling for more nuanced mixed methods approaches to capture the complexities of each individual case. In what follows, I present one possible approach to studying discursive content from SNSs.
Methodology

Data collection [13]

Using Facebook Graph API, I collected the maximum allowed number of posts [14], and their respective comments from 32 Facebook pages that were active in reporting and engaging with political, and later military, events that have taken place in Syria since 2011 [15]. I selected these pages based on the minimum requirement of having at least 15,000 [16] followers at the time of the initial collection in 2016 and were supportive of either the regime or the opposition, based on their “about” sections at the time, with the exception of YawmiyatKazekeh and vcdc. The former covered people’s daily lives, including the ongoing war (66 percent of their posts were not war-related) but the posts that were war-related indirectly support the regime, and the latter was a human rights NGO that documented war deaths, kidnappings, and forced disappearances. vcdc could be considered the only non-affiliated page even though their leading administrators were activists in the revolution [17]. However, vcdc has the least number of posts and related comments. A detailed description of the numbers, posts, and comments per page is provided in Table 1. I included 18 active pro-regime pages and 12 pro-opposition pages because the total number of comments on the pro-regime pages was much lower than on pro-opposition pages (see Table 2).

<table>
<thead>
<tr>
<th>Political Support</th>
<th>Pages</th>
<th>Posts</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pro-regime</td>
<td>18</td>
<td>41,901</td>
<td>1,432,477</td>
</tr>
<tr>
<td>Neutral</td>
<td>2</td>
<td>4,127</td>
<td>1,297,947</td>
</tr>
<tr>
<td>Pro-opposition</td>
<td>12</td>
<td>32,478</td>
<td>2,446,577</td>
</tr>
<tr>
<td>Total</td>
<td>32</td>
<td>78,506</td>
<td>5,177,001</td>
</tr>
</tbody>
</table>

Table 2: Totals for examined Facebook pages.

Even though most of these pages were created after 2011, I started data collection on 1 January 2011 and ended it on 31 May 2018, in order to capture changes in commenting behavior, if any, throughout the seven years of unrest in Syria and following major military and political events. The total number of unique posts was 78,506 and the total number of comments was 5,177,001. All these pages used Arabic as the main language in their posts and the majority of comments were also written in Arabic. The author is also a native Arabic speaker who analyzed, translated, and transcribed all of the Arabic text in this paper.

Machine learning and content analysis

While Ben David and Fernández used frequency and co-occurrence textual analysis to identify main topics generated from posts and comments of each political party page, that approach captures thematic clusters but does not capture relations between each post and comments which the proposed method used in this work allows.

This paper is part of a larger study in which I used a supervised machine-learning approach to classify the full dataset of posts and their respective comments. Initially, I manually classified a sample of 3,000 randomly chosen unique comments and their respective posts. To prevent overfitting of the model my software started with randomly choosing 2,700 comments as a training set and 300 comments as a testing set. The training set was used to teach the machine how to classify, based on manual classification conducted on 2,700 comments and their respective posts. The testing set was used to test how well the
machine performed by automatically classifying 300 comments and their respective posts and compared the results to the manual classification initially conducted. The classification process included multiple coding layers in order to capture complex meanings expressed in the posts and their respective comments. The first layer coded the post and its respective comment as pro-regime, pro-opposition, or neutral. The second layer coded them as religious/non-religious, sectarian/non-sectarian, violent/non-violent, curse word/no curse words, reconciliation/non-reconciliation.

For the purpose of this study, I only used the first layer of classification and the religious part of the second layer; I only analyzed the political affiliation of the comments and their respective posts. For example, "اللهم بشار بشارواتباع خنزير هل انصرناعلى اللهم المسلمين واعز اسام انصر" which translates to “O Allah, support Islam and the most cherished Muslims, O Allah support us against Bashar the pig and the followers of Bashar the mule” was classified as pro-opposition. And "الله يحميك يا فخر سوريا الأسد ويهمي أولادك والسيدة الأولى زهرة الياسمين" translates to “May God protect you, Syria’s pride, and protect your children and the first lady, Jasmine flower.” This comment was classified as pro-regime. “اللهم يستر” which translates to “God prevent the worst,” was classified as neutral. These examples are easy to classify, as they included indicators of their positions. However, other posts and comments with no stated positions were more difficult to classify, I include below a detailed discussion for ways to overcome these difficulties.

**Methodological issues studying SNSs in Arabic**

The fields of computer science and computational social sciences have developed multiple approaches to conducting text analysis. Depending on the questions asked, researchers might use topic modelling approaches to extract topics from text data, such as topics covered in news articles (Newman, *et al.*, 2006), recommendations of journal articles (Wang and Blei, 2011), or even SNSs analysis (Hong and Davison, 2010). Researchers might also use sentiment analysis to classify users’ opinions and sentiments (Liu, 2012; Fang and Zhan, 2015), usually classified as negative, positive, or neutral. Although there have been many attempts to write sentiment analysis software for the Arabic language, Boudad, *et al.* (2017) noted that Arabic is one of the most challenging languages to conduct this sort kind of analysis. They rightly argued that the three main varieties of Arabic — Classical Arabic (CA) [19], Modern Standard Arabic (MSA) [20], or Dialectical Arabic (DA) [21] — create the initial obstacle for such studies of Arabic text online. In previous studies of Arabic sentiment analysis, three approaches — supervised, unsupervised, and hybrid — have been used for studying and classifying a wide range of text such as tweets, Facebook comments, blogs, book reviews, and product reviews.

My dataset of the collected Facebook posts and comments included the three varieties of Arabic language mentioned above (CA, MSA, DA), with multiple dialects and sometimes a combination of Arabic and English, French, Spanish, or Portuguese. The Arabic in the comments reflected non-Syrian regional involvement in the Syrian war, Syrians in the diaspora, as well as a range of dialects within Syria. These variables made it difficult to use the unsupervised or hybrid approach because these approaches relied on a predefined lexicon of weighted words, which created multiple problems due to the non-unified Arabic text (CA, MSA, DA) in this dataset as well as the lack of lexicon datasets such as Natural Language Toolkit (NLTK) for English. In order to address these difficulties, I chose a supervised, machine-learning approach using a deep neural network (DNN) classifier with embedding, detailed in the section later.

This approach transformed words into unique numbers regardless of the dialect, or language used. Then, these words became sequences of numbers that formed sentences, and in order to capture the nuances between these sentences, embeddings assigned multidimensional distances between words based on these sequences. The classifications have a weight that was calculated in relation to the sequence of words as a whole. The model will be able to identify a text’s classification based on words and their distances to each other in the sequence in order to predict a score for a text, in this case whether it was supportive or opposing of the Syrian government and whether the text was a religious expression.
Data classification

Pre-processing

In selecting content for this study, I excluded all Facebook posts and their respective comments that only contained image or video content with no meaningful text description. I included comments and posts that used a mix of Arabic and other languages. I applied a pre-processing filter to exclude symbols, with the exception of emojis. Then, I created a bag of words out of the total number of posts and comments and ranked them based on the total word count in the dataset. Figure 1 shows a word cloud for the top 10,000 words from the bag of post words. Figure 2 illustrates a word cloud for the top 10,000 words from the bag of comment words.

Figure 1: Word cloud visualization of all Facebook posts in the original Arabic.
I created an embedding file from all the words from both sets with a frequency greater than 10. The bag of words was also used as a starting point for the content analysis that started with the most frequent words in order to identify recurring expressions, if any. Table 3 shows the top 10 words found in the posts and the top 10 words found in the comments. In the comments section, four of the top 10 words formed the following phrase: حسبنا الله ونعم الوكيل، literally translated to “Sufficient for us (is) Allah and (He is the) best, [the] Disposer of affairs.” This leads to the conclusion that a sizeable portion of the comments included this phrase, a point I return to later. The total number of unique words was 29,510 from the posts and 175,195 from the comments.
The Syrian War on Facebook: A call for mixed methods

**Table 3:** Comparison of top ten words found in posts versus comments.

<table>
<thead>
<tr>
<th>Posts</th>
<th>Frequency</th>
<th>Word</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>17436</td>
<td>سورية</td>
<td>Syria</td>
<td></td>
</tr>
<tr>
<td>15686</td>
<td>دمشق</td>
<td>Damascus</td>
<td></td>
</tr>
<tr>
<td>15172</td>
<td>الرئاسي</td>
<td>The Syrian</td>
<td></td>
</tr>
<tr>
<td>14071</td>
<td>سوريا</td>
<td>Syria</td>
<td></td>
</tr>
<tr>
<td>11533</td>
<td>الأسد</td>
<td>Assad</td>
<td></td>
</tr>
<tr>
<td>10647</td>
<td>الجيش</td>
<td>Army</td>
<td></td>
</tr>
<tr>
<td>10001</td>
<td>الله</td>
<td>Allah</td>
<td></td>
</tr>
<tr>
<td>9448</td>
<td>السوري</td>
<td>The Syrian</td>
<td></td>
</tr>
<tr>
<td>9250</td>
<td>اليوم</td>
<td>Today</td>
<td></td>
</tr>
<tr>
<td>8324</td>
<td>مدينة</td>
<td>City</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comments</th>
<th>Frequency</th>
<th>Word</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2257630</td>
<td>الله</td>
<td>Allah</td>
<td></td>
</tr>
<tr>
<td>411068</td>
<td>الوكيل</td>
<td>agent</td>
<td></td>
</tr>
<tr>
<td>400361</td>
<td>واعظ</td>
<td>and best</td>
<td></td>
</tr>
<tr>
<td>351444</td>
<td>الهم</td>
<td>May Allah</td>
<td></td>
</tr>
<tr>
<td>307770</td>
<td>بارب</td>
<td>Oh Allah</td>
<td></td>
</tr>
<tr>
<td>270616</td>
<td>حبين</td>
<td>Sufficient is</td>
<td></td>
</tr>
<tr>
<td>236719</td>
<td>سوريا</td>
<td>Syria</td>
<td></td>
</tr>
<tr>
<td>165538</td>
<td>والله</td>
<td>I swear</td>
<td></td>
</tr>
<tr>
<td>125848</td>
<td>شبار</td>
<td>Bashar</td>
<td></td>
</tr>
<tr>
<td>120868</td>
<td>يحميك</td>
<td>May He protect you</td>
<td></td>
</tr>
</tbody>
</table>

**Manual classification**

The manual classification process started with a random sample of 3,000 comments and their respective posts. Two separate manual classifications were implemented in order to limit the author’s bias and personal point of view. One classification was implemented by the author and the other was implemented by an undergraduate student research assistant who was also a native Arabic speaker. I developed an interface that allowed a less tedious manual classification process. This interface made the comparison process of the separate manual classifications easier during weekly check-in meetings with the research assistant. The goal of these meetings was to discuss the reasons behind differences in classifications by the author and research assistant, when they occurred, which were automatically flagged. We went through these comments and discussed reasons for different classifications and then the author decided which classification to adopt for the training process of the machine learning algorithm. These discrepancies decreased as we completed more classifications.

The posts were classified in the following manner: peaceful, reconciliatory, religious, sectarian, support (Regime/Neutral/Opposition), type (open), report death, war-related, violent, opposition faction (open), anti-Hizballah, anti-Iran, anti-Russia, anti-USA, anti-Saudi, anti-Europe. The comments were classified based on these categories: peace, reconciliatory, religious, sectarian, support, sarcastic, irrelevant, violent, type, opposition faction, anti-Hizballah, anti-Iran, anti-Russia, anti-USA, anti-Saudi, anti-Europe. Here is an example of a post from a page that supported the opposition. The text included in the post was translated as:

As if the apocalypse is happening ... The first moments of targeting the National Hospital in the city of #Maarat_Anuman more than eight Russian military raids tonight. Newborn babies were injured in their nurseries, and the wounded who were taken from the Kafir Nabeel massacre and Khan al-Sibel were targeted again inside the hospital! The reporter of the Syrian Revolution Network # Muath_Shami, in a report summarizes all the talk. https://youtu.be/fRHSrvXJLDs

I categorized this post as reporting war, political, report death, supports opposition, and as anti-Russia. As for the comment on this post which translates to “Sufficient for us (is) Allah and (He is the) best, [the] Disposer of affairs and there is no power but from God the Almighty” is categorized as religious and supports the opposition.
DNN with embedding

A DNN is an artificial neural network with several layers between the input and output layers of data. Each level “learns” to transform the input data into a more abstract representation, the “deep” in “deep learning” refers to the number of transformations through which the data is analyzed. The benefit of using a DNN is the ability to move away from binary representations for a wider view of the text used in Facebook posts and comments. Given that the content I analyze in this paper is mainly in Arabic, for which no large dependent textual representation exists (even the Natural Language Toolkit is not applicable here as the Arabic NLTK cannot combine all three genres of Arabic), using a DNN makes the analysis possible. In addition, the fact of using a DNN with embedding allows for the representations of words with varying degrees of similarity, enabling a faster and more accurate analysis. In short, this study involves analyzing millions of textual interactions, and benefits from the ability to increase the accuracy of the model of representation of words, especially in Arabic [22].

Simply put, the embeddings are words with specific coordinates plotted on a two-dimensional axis. In this case, support of the Syrian government is plotted on one extreme and support of the opposition on the other. So, words that only show up in comments that support either the regime or the opposition will take the extreme coordinates and words that only show up in comments that are neutral show up in the middle. Embeddings do not stop at this level; embeddings are plotted in a multidimensional space (see Figure 3) which takes into account the sequences of words and plots in order to determine support. For example, the word ‘Assad’ might be found in both comments of support and opposition, so Assad is plotted in relation to the sequence of words based on the training dataset (the data that was manually classified).

Figure 3: Example of the word ‘respondent’ with the closest 100 potential possibility for plotting.
Using a DNN classifier and the embedding file that I produced, I classified the posts and their respective comments in terms of political support as pro-regime, pro-opposition, and neutral, in order to measure a correlation between the political affiliations of the pages, support of the posts as well as comments, to check whether comments from non-supportive users appeared in the comments section [23]. The classification process was divided into two stages: first, the DNN classifier categorized the posts originated by the pages, and second, it classified the comments to the respective posts. The DNN classifier achieved a 92.33 percent accuracy rate when classifying the 300 posts testing set, as illustrated in the confusion matrix in Figure 4.

![Confusion Matrix for Posts (With Support)](image_url)

**Figure 4:** Confusion matrix for posts that express support in the testing set.

Note: Out of the 38 posts that expressed support towards the regime, 36 were correctly categorized (2 were categorized as neutral and 0 as support towards the opposition); out of the 114 posts that were neutral in their expression of support, 107 were correctly categorized; and out of the 137 posts that expressed support towards the opposition, 134 were correctly categorized.

My DNN classifier failed to classify the political support of the comment’s text using comment_text as the single input variable to analyze and classify. The accuracy rate was 48.2 percent.
I attribute this failure to the similarity of the language and narrative used by both the pro-regime and pro-opposition. For example, many people used "الْوَكِيلُ وَنِعْمَهُ هُّ حَسْبُنَا, literally translated to “Sufficient for us (is) Allah and (He is the) best, [the] Disposer of affairs” to support the post’s position. Without knowing the political position of the author of the post, one cannot predict whether the comment supports or opposes the post’s content (which is usually implicitly or explicitly supporting one side or the other). When adding the page’s political affiliation and the post’s reflecting political side features (support) as input variables along with comment_text, this model achieved a rate of around 92 percent accuracy (see Figure 5).

**Figure 5:** Confusion matrix for comments that express support in the testing set.
Note: Out of the 70 comments that expressed support towards the regime, 64 were correctly categorized (1 was categorized as neutral and 5 as support towards the opposition); out of the 74 comments that were neutral in their expression of support, 65 were correctly categorized;
and out of the 156 comments that expressed support towards the opposition, 146 were correctly categorized.

One might think that the failure could be attributed to the model of classification itself but the same algorithm with the same training and testing sets was able to achieve a rate of around 90 percent accuracy (see Figure 6) when classifying whether the sentences were religious or not religious based on the text only regardless of the association of the page or post.

**Figure 6:** Confusion matrix for comments classified as religious/non-religious in the testing set.
Note: Out of the 64 comments that were not religious, 51 were correctly categorized; and out of the 236 comments that were religious, 221 were correctly categorized.
Figure 7: Pie chart for comments on pro-regime pages.

Figure 8: Pie chart for comments on pro-opposition pages.
Additionally, the phrase “the burning of Muslims in Burma,” which was among the dataset, received a prediction score of 0.035 (where 0 for non-religious and 1 for religious) even though it has the word Muslim in it. This confirms that the analysis did not rely on the meaning of words, rather it depends on a sequences of words. In sum, the analysis correctly determined whether a post was an expression of support or opposition, but contextual information of a page and posts’ political affiliations was needed to determine what party, whether regime or opposition, towards which the expression was directed.

Finally, using the model generated above, I classified all of the posts and their respective comments from the dataset.

**Results**

The automatic classification of posts and their respective comments led to the following results: 76.81 percent of the posts on pro-regime pages supported the regime and 23.19 percent were neutral, meaning they did not relate to the war. These neutral posts included, for example, coverage of a TV series during the religious month of Ramadan, or the work of some ministries such as those working to fix power outages caused by the war, economic and social initiatives, or even miscellaneous news unrelated to Syria. On the pro-opposition pages, 86.58 percent of the posts were directly related to the war and 7.42 percent of the posts were not relate to the war. Examples of the latter included reporting that “the borders with Turkey are open during the month of Ramadan.” As for the comments on these posts, on the pro-regime side, 96 percent of the comments supported the regime, two percent supported the opposition, and two percent were neutral (see Figure 7). On the pro-opposition side, 91 percent of comments on the posts (relevant to the war) were supportive of the opposition and/or were anti-regime, while seven percent were supportive of the regime, and two percent were neutral (see Figure 8).

**God is everywhere**

The word “Allah” appeared in 1,233,394 comments on opposition pages, which means “Allah” was included in 50.4 percent of the total number of comments on opposition pages. “Allah” also appeared in 554,410 comments on regime pages and 490,254 of neutral pages, which means it was respectively included in 38.7 percent and 37.7 percent of the total comments of the later pages. More specifically, religious comments that include “Allah” were distributed as follows: 72 percent for the opposition, 51 percent for the regime, and 43 percent for the neutral pages. This reflects the fact that religious discourse was used more on opposition sites than on neutral ones, but it also reflects the different feeling states of people on various sides of the conflict. Most religious terminologies overlapped on all sides of the conflict, making it difficult to distinguish which side each term supported, but religious terminologies that reflected helplessness and despair were overwhelmingly used by the opposition. For example, the expressions “Sufficient for us (is) Allah and (He is the) best, [the] Disposer of affairs” and “There is no might nor power except in Allah” were overwhelmingly used by commentators to the opposition pages (91,346 and 58,714 respectively), followed by commentators on neutral pages, (16,757 and 9,855 respectively), and finally by those on regime pages (2,009 and 1,088 respectively). I address the implications of these numbers in the next section.

**Discussion**

The results of this study, shown in Figures 7 and 8, confirmed that these Facebook pages were producing echo chambers. This supports the findings of previous studies that confirm that SNSs, in general, produce
The results of this study confirm that the war was not translated to Facebook pages as opposing parties were not fighting each other on these spaces. The fact is these pages were producing segregated venues where like-minded people were posting their comments. These pages were not playing the role of public spheres as argued by Al-Mustafá as they were not producing venues where possible conversations and/or discussions were taking place. A noticeable finding that also deserves future study and analysis is the absence of neutral voices from either space.

The word “Allah,” “God” appeared in more than half of the total number of comments and was used in many variations, including religious (in verses, sayings, praises) and cultural (in figures of speech). The religious uses varied between commentators on different pages, and commentators on opposition pages used more religious comments than those on the neutral and regime pages. But this finding begs a more in-depth content analysis of the posts published, a study of the pages and whether they differed by type of content they publish and that was a cause for the greater incidence of religious comments on these pages. Another important finding was that use of statements that expressed despair and helplessness were used primarily by commentators on opposition pages, a finding which indicates that a longitudinal study of the use of these statements could reveal whether statements of despair were related to the content published on the pages or to the unfolding of the events on the ground. For example, were these comments used more frequently after a battle loss? Or did the published content of opposition pages stimulate these feelings?

Finally, it is intriguing that commentators on neutral pages used more statements that revealed helplessness and despair than did commentators on regime pages, which could indicate that regime supporters were expressing helplessness and despair on pages that were not directly linked to the regime. One might even conclude that the public aspect of the data on these pages was preventing users from commenting online due to the volatile political, military, and economic conditions that people were living in. This was likely to be due to the decades of authoritarian censorship and dissimulation, as Lisa Wedeen argued:

> People [in Syria] are habituated to provide publicly acceptable responses but not necessarily through outright lies. They use their imaginations to invent modes of evasion; they find ways of both speaking truth and yet reiterating the regime’s idealized presentation of itself. [24]

**Discussion in relation to ethnographic fieldwork**

The 18 months of ethnographic fieldwork I conducted in 2016–2019 in Lebanon and Syria included 82 interviews, and many informal discussions and field observations, and my findings reflect a conclusion...
Eighty-one out of 82 interviewees noted that, regardless of the outcome of the war, Syrians will reconcile and will live together in their country. For example, Ahmed interviewed in Lebanon: a young professional from Damascus who was active in the 2011 uprisings and was jailed twice, tortured, and had to flee to Lebanon repeatedly said in the interview that Syrians will live together after the war. “At the beginning after I came here [Lebanon] I did not want to see any supporter of the regime, I felt like they were the ones who tortured me. But look now A’lia is the one who put you in touch with me and we are real friends now. Syrians have lived together and will eventually be living together after the war. We don’t have a choice. We are good people by nature.” A’lia, who I also interviewed in Lebanon, a young professional working in Beirut and a staunch supporter of the regime, was the one to put me in touch with Ahmed knowing his opposite political views. A’lia also emphasized the peaceful past in Syria. She also stated that Syrians lived together before 2011 and will live together after the war.

Others like Hisham, who was also an opposition activist in Homs, kept repeating that “you will see, as soon as the regime falls and Assad is gone, Syrians will realize that they only have each other and no one can eliminate the other.” A’mer, who supports the regime, similarly stated that “as soon as ISIS and the other terrorists are gone, Syrians will realize that the bloodshed was not worth it, and we will get back together. It might take time, but eventually we are all Syrians.” Here, as in many other interviews the process of how to get back together and live after all the bloodshed, destruction, and suffering seems to be just wishful thinking. But a youth group I studied in Lebanon started this conversation with periodic meetings, where they discuss, argue, and debate the ongoing military and political developments on a regular basis. These debates became heated and contested. For example, Amir stormed out of the meeting after a heated argument with Rasha but he returned to the next meeting and was able to continue the discussion with Rasha. The members of this group are able to find a place for discussions even with their stated opposing political alignments. Such spaces, where potential reconciliation is taking place, cannot be deduced from an online analysis.

My interviews and participant observations also reveal that regime supporters are now openly criticizing the regime, even the president and the international allies, while the Syrian army is still above all criticism. This criticism is almost absent online.

Limitations and future work

My current study presents multiple limitations. Some are related to large-scale data analysis in general, and others are related to the size of the training dataset. In order to achieve a higher accuracy rate in the full classification model, I will, in future work, increase the size of the training set to 10,000 classified documents and then retest my model. These limitations of supervised, machine-learning are to be expected, as manually classifying and cross-checking these classifications is time-consuming and requires ample resources. Future studies will also address the absence, in this study, of distinguishing between the multiple factions of the opposition, by focusing on conversation changes on opposition pages. Other limitations are specific to the online data for example, one can conclude from this analysis of Facebook pages that Syrians suffer from an irreconcilable divide and that they are not going to reconcile as a people in the foreseeable future.

Conclusion

In contrast to the majority of studies about social network sites, which usually focus on the categorization
of users’ political positions, this study provides an example of another way to examine conversations on public Facebook pages concerning the Syrian civil war. This paper differs from others in its approach in that it is not producing a social network for users across multiple Facebook pages, but rather is classifying posts and comments as text entities regardless of the users who publish them. This approach can reveal changes in the dynamics of the conversations taking place on these pages. The initial results of this study are promising for a textual analysis that moves beyond a labeling or a limited use of large datasets, and points to potential analyses with multi-layered results. The paper concludes that online conversations alone cannot give scholars a full sense of people’s opinions, especially when compared with off-line fieldwork.

About the author

Toni Rouhana is a Postdoctoral Research Associate in the Civil War Paths Project, part of the Department of Politics and International Relations at the University of Sheffield.
E-mail: t [dot] rouhana [at] sheffield [dot] ac [dot] uk

Acknowledgments

This work was supported by the Global Religion Research Initiative at the University of Notre Dame.

Notes


2. Most of these security branches were either set up by the French colonial rule or were offshoots of the system set in place by colonial France.

3. Even the O’Callaghan, et al. (2014) study about Syria that argued for a different approach ended up with a social network showing the political views of the sampled accounts that they studied.

4. With a few exceptions (Wolfsfeld, et al., 2013) most authors argue that social media use increases as a consequence of increased political action. Rather than assuming that an increase of social media use led to an increase in political activity, Wolfsfeld, et al. demonstrated the opposite.

5. Howard, et al., 2011, p. 3.

6. The social network analysis included a final set of 652 Twitter accounts and 295 unique YouTube channels.


10. These conclusions are questionable because Al-Mustafa did not provide sufficient evidence for the causal relationship between online discussions and deviations that took place on the ground. But this does not reduce the importance of this study.

11. A full Internal Review Board protocol was developed and approved by the Office of Research at UC
Santa Cruz to conduct this research.


13. No IRB was necessary as these pages are public and I was not singling out specific comments or posts that might point to users’ specific virtual identities.

14. Graph API allows for approximately 600 posts per year and a maximum of 25,000 comments per post to be collected. I only had two posts that had more than 25,000 comments which exceeds the limits that Facebook allows in terms of collected comments, so I collected the top 25,000 comments for each of these posts.

15. Some of these pages have been taken down since the end of data collection in 2018 and many posts and comments are not accessible anymore in light of Facebook’s recent data access policies.

16. All the Facebook pages that had less than 15,000 followers did not have enough comments in order to be studied. For example, vdc.sy which had 15,000 followers at the time and therefore included in this study only has 1,010 comments (see Table 1).

17. Razan Zaitouneh, the head of the Violations Documentation Center in Syria (VDC) in 2013, was vocal in opposing the regime and the atrocities of some opposition groups was arrested by the regime (https://www.theguardian.com/world/2011/may/21/syria-women-unrest-repression) and later abducted by an unknown group some claimed was part of the opposition (http://vdc-sy.net/our-story/). But the NGO documented casualties and human rights violations from all sides of the conflict.

18. They have an outstanding literature review of research of sentiment analysis in Arabic.


20. The language used in formal communication.

21. There are thousands of dialectical Arabic variations.

22. For an overview of DNNs, see Schmidhuber (2015). For more on embedding, see Mikolov, et al. (2013).

23. Due to the data being in Arabic, and the limited possibilities for multi-labeling in Arabic at the time of classification, I decided on using DNN with embeddings and did not use any other classification algorithms to cross-validate.


References


The Syrian War on Facebook: A call for mixed methods


The Syrian War on Facebook: A call for mixed methods


The Syrian War on Facebook: A call for mixed methods
by Toni Rouhana.
*First Monday*, volume 27, number 9 (September 2022).
doi: [https://dx.doi.org/10.5210/fm.v27i9.11557](https://dx.doi.org/10.5210/fm.v27i9.11557)