

# Misogynistic Disinformation on Arabic Twitter: A Case Study of Online Harassment Against Women Journalists in the Middle East

Authors: *Marc Owen Jones and Wajdi Zaghouani*

Affiliation: *Hamad Bin Khalifa University*

## Introduction

We published a case study in 2021 titled 'State-aligned misogynistic disinformation on Arabic Twitter: The attempted silencing of an Al Jazeera journalist' (Jones 2021). This followed an incident that occurred in June of 2020, where Ghada Oueiss, a principal anchor of the news channel Al Jazeera Arabic, received online harassment that started with her phone being hacked anonymously and her private photos shared online. In addition to hacking and sharing Oueiss's private information, false offensive narratives were also spread. Some of these tweets had been retweeted over 40,000 times (Oueiss, 2020). Another Al Jazeera anchor, Ola Al Fares, was also targeted to a lesser extent. This hateful campaign was mainly trending in Saudi Arabia and the Gulf region, and it was a large-scale attack that spread very quickly. This paper helps to fill the gap on disinformation studies in the Middle East.

Recent research has focused on various aspects of hate speech, sentiment analysis, and political discourse in Arabic social media. Charfi et al. (2024) introduced MARASTA, a multi-dialectal Arabic cross-domain stance corpus, while Laabar and Zaghouani (2024) provided multi-dimensional insights into stance, sentiment, and emotion in Facebook comments related to Tunisia's July 25 measures. AlEmadi and Zaghouani (2024) explored the emotional toll and coping strategies of annotating hate speech data. Magnossão de Paula et al. (2023) proposed a solution for hate speech detection in Arabic languages using transformers and ensemble methods. Shestakov and Zaghouani (2024) analyzed conflict through data by creating a dataset on the digital framing of Sheikh Jarrah evictions. Shurafa, Darwish, and Zaghouani (2020) studied political framing in the context of the US COVID-19 blame game. Zaghouani, Mubarak, and Biswas (2024) built a multi-label hate speech annotated Arabic dataset, and De Paula et al. (2022) participated in the Arabic Hate Speech 2022 shared task, focusing on offensive language and hate speech detection using transformers and ensemble models.

When women are targeted in ways that paint them as lacking intelligence, being untrustworthy, or driven by emotions and desires, their public image is tainted. Violence against women often includes disinformation, a central element to hate speech, with the goal of discrediting the victim. Since the Middle East rates highly on gender inequality, and women's rights movements often find

themselves the subject of harassment, women journalists who get targeted are unlikely to find support from the state.

All around the globe, those working in the media are often quickly targeted when political disagreements arise. Women journalists and activists continue to face online misogyny and threats, and there has been an increase in misogynistic attacks and harassment of women on social media during the COVID-19 pandemic (Jazeera, 2020). Notably, the online attack on Ghada Oueiss happened at a sensitive time for Arab journalists, since it was likely motivated by Oueiss's outspoken criticism of Saudi Arabia following the assassination of the Saudi journalist Jamal Khashoggi by the Saudi State. In the context of the 2017-2021 Gulf Crisis, which saw the isolation of Qatar by KSA, Egypt, Bahrain and UAE, this was, therefore, considered a broader attack from UAE and Saudi Arabia at Al Jazeera.

It is evident that individuals who share misogynistic content online, including offensive posts directed at prominent women, do not necessarily face any consequences. Women on social media then feel that they cannot express their opinions freely, despite their desire to change the narrative and speak against certain social norms. Women can also become unlikely to pursue political roles. It does not help that populism has been on the rise and we see more authoritarian leaders who wish to impose patriarchal gender roles, encouraging gendered violence. Furthermore, with the accessibility of the internet, populism and misogyny can easily travel across borders. In addition to right-wing populism seen in the US and Europe, some of the blame also falls on the social media companies that are not monitoring and censoring gendered abuse, especially when it occurs in languages other than English.

## **Methodology**

We examined how three tactics of online harassment, shaming, discrediting, and intimidation (Sobieraj, 2017) were used against Oueiss in a Twitter dogpile attack. We analyzed the online campaign against Ghada Oueiss over the five-day period in June 2020 and highlighted the drivers of the campaign. We explored memes, cartoons and other content that can be missed by online platforms' censoring attempts that are keyword-based. We examined the scale, speed, and intensity of the campaign as well as its outbreak outside of social media. The search criteria for the tweets to be downloaded and analyzed were three hashtags related to the campaign that were trending in Saudi Arabia and were directly concerning the victims. A total of 24,713 tweets, retweets, replies, or mentions were downloaded from Twitter, all produced by 10,410 different accounts. The analysis pointed to the most influential accounts and the most retweeted tweets in the 2020 attack, in addition to exploring other attacks that contained sexualized or demeaning language that Oueiss had been a victim of between 2014 and 2020.

## **Results**

A scraping of tweets with the three hashtags returned 11,000 tweets between 5 June 2020 and 7 September 2021. Network analysis of the 24,713 interactions over the five days in June 2020 showed the following: the two highest reported locations for accounts using the hashtags were Saudi Arabia (86%) and United Arab Emirates (UAE; 6%). The most retweeted tweets came from pro-Saudi regime accounts. In addition, accounts using the hashtags had a history of spreading disinformation. Some of the most popular interactions were sexist and otherwise offensive jokes, which could and were perceived as threats by victims. Other interactions had other goals, including provoking ethnic tension by suggesting that Qatari citizens should be appalled by the foreign journalist's behaviour. Other interactions were made up of derogatory cartoons, attempted to paint Qatar and those who work in Qatar as sexually deviant, insinuated violence, or consisted of manipulated and fake tweets attributed to Oueiss.

The attack that Oueiss and Al Fares received traveled across multiple social media platforms, broke out beyond social media, and was amplified by celebrities and political figures who endorsed it. Non-anonymous accounts also attacked Oueiss, and she reported having blocked over 9,000 accounts. CNN Arabic (2020) produced a piece that portrayed the situation as a petty quarrel between the two women journalists, adding to the stereotype that women are inherently flawed with qualities such as jealousy and vindictiveness.

We analyzed and reported the most used offensive and misogynistic terms in the interactions that were part of the campaign. The data showed a spike in gendered abuse towards Oueiss in 2018 around the time that Jamal Khashoggi had been assassinated. This was due to many Saudis defending the Saudi regime, as the Crown Prince was accused of playing a role in the murder. The other significant spike in the use of the misogynistic terms was in June 2020, at the time of the anonymous hacking and the abuse that followed.

## **Discussion**

We discussed the way that social media attacks tend to show certain trends in their mechanism and the stages and sequence they follow, which makes them appear calculated (Edson et al., 2021). Online attacks, which are psychological, can become physical attacks, and because online attacks could be orchestrated, they are perceived as more threatening, as especially when those engaging in them are high level or well-connected officials. The campaign on Oueiss was not an isolated incident, and journalists continue to have their privacy breached and their safety threatened. Therefore, we call for the importance of immediate intervention when online abuse like this occurs, as an analysis that allows understanding of what could ensue would help prevent the abuse from circulating and becoming unmanageable. The sooner it is dealt with, the less likely it is to cause as much harm. In addition, there is a need for more understanding of online misogyny in the context of the Arab region to best support those in positions that might attract cyberhate. We

concluded by asking whether the outbreak of social media hate campaigns into other outlets is strategic and deliberate. We also highlight the importance of immediately and effectively responding to this type of incident and wonders about the true role of states in spreading and encouraging disinformation and mal-information. In addition, we pointed out the possibility for victims of online hate to hold the attackers accountable, but that this does not come without its own downfalls especially in states that do not punish online abuse. Finally, although Twitter and other social media platforms have harassment policies, it does not seem that they enforce them globally, but rather only in the Global North.

## **Acknowledgements**

This work was made possible by grant NPRP13S-0206-200281 from the Qatar National Research Fund (QNRF) Part of Qatar Research Development and Innovation Council (QRDI). The contents herein reflect the work and are solely the authors' responsibility.

## **Reference**

CNN Arabic (2020, June 10). Saudi Qatar Social Ghada Ola-Abdulatif Reactions. <https://arabic.cnn.com/middle-east/article/2020/06/10/saudi-qatar-social-ghada-ola-abdulatif-reactions>

Charfi, A., Ben-Sghaier, M., Atalla, A., Akasheh, R., Al-Emadi, S., & Zaghouani, W. (2024). MARASTA: A multi-dialectal Arabic cross-domain stance corpus. In Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)\*.

De Paula, A. F. M., Rosso, P., Bensalem, I., & Zaghouani, W. (2022). Upv at the arabic hate speech 2022 shared task: Offensive language and hate speech detection using transformers and ensemble models. Proceedings of the 5th Workshop on Open-Source Arabic Corpora and Processing Tools with Shared Tasks on Qur'an QA and Fine-Grained Hate Speech Detection, 181–185.

Edson C. Tandoc, Karryl Kim Sagun & Katrina Paola Alvarez (2021): The Digitization of Harassment: Women Journalists' Experiences with Online Harassment in the Philippines, Journalism Practice, DOI: 10.1080/17512786.2021.1981774

Jazeera, A. (2020, November 11). Digital misogyny: Online abuse of women surges during COVID. Aljazeera.com; Al Jazeera. <https://www.aljazeera.com/economy/2020/11/11/women-bear-brunt-of-digital-abuse-ascovid-drives-people-online>

Jones, Marc Owen. "State-aligned misogynistic disinformation on Arabic Twitter: The attempted silencing of an Al Jazeera journalist" *Open Information Science*, vol. 5, no. 1, 2021, pp. 278-297. <https://doi.org/10.1515/opis-2020-0126>

Laabar, S., & Zaghouni, W. (2024). Multi-dimensional insights: Annotated dataset of stance, sentiment, and emotion in Facebook comments on Tunisia's July 25 measures. In *Proceedings of the Second Workshop on Natural Language Processing for Political Sciences co-located with the 2024 International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*.

Magnossão de Paula, A. F., Bensalem, I., Rosso, P., & Zaghouni, W. (2023). Transformers and Ensemble methods: A solution for Hate Speech Detection in Arabic languages. *ArXiv E-Prints*, arXiv-2303.

Maryam M. AlEmadi and Wajdi Zaghouni (2024). Emotional Toll and Coping Strategies: Navigating the Effects of Annotating Hate Speech Data. In *Proceedings of the Legal and Ethical Issues in Human Language Technologies (LEGAL2024) co-located with the 2024 International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*.

Oueiss, G. (2020, July 8). I'm a female journalist in the Middle East. I won't be silenced by online attacks. *The WashingtonPost*. <https://www.washingtonpost.com/opinions/2020/07/08/imfemale-journalist-middle-east-i-wont-be-silenced-byonline-attacks/>

Shestakov, A., & Zaghouni, W. (2024). Analyzing conflict through data: A dataset on the digital framing of Sheikh Jarrah evictions. In *Proceedings of the Second Workshop on Natural Language Processing for Political Sciences\* co-located with the 2024 International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*.

Shurafa, C., Darwish, K., & Zaghouni, W. (2020). Political framing: US COVID19 blame game. *Social Informatics: 12th International Conference, SocInfo 2020, Pisa, Italy, October 6–9, 2020, Proceedings 12*, 333–351.

Sobieraj, S. (2017). Bitch, slut, skank, cunt: Patterned resistance to women's visibility in digital publics. *Information, Communication & Society*,21(11), 1700–1714. <https://doi.org/10.1080/1369118x.2017.1348535>

Zaghouni, W., Mubarak, H., & Biswas, M. R. (2024). So hateful! Building a multi-label hate speech annotated Arabic dataset. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)\**.