DEMOCRACY
REPORTING
INTERNATIONAL

NOVEMBER 2022 - APRIL 2023

ONLINE PUBLIC DISCOURSE IN THE MENA REGION

PERSISTENT TACTICS OF DISINFORMATION AND AN INCREASE IN ONLINE GENDER-BASED VIOLENCE

Warning:

Social media monitoring reports contain potentially disturbing content that may be distressing for some readers.

Democracy Reporting International is sharing this content only for scientific and research purposes.

This third report has been produced by DRI and its partners for the project "Words Matter". The report covers the period from November 2022 - April 2023:

الكلمة تفرق
Words Matter

DRI Partners

RASED راصد

مركز الحياة
AL HAYAT CENTER

مراقبون
Mourakiboun

مهارات

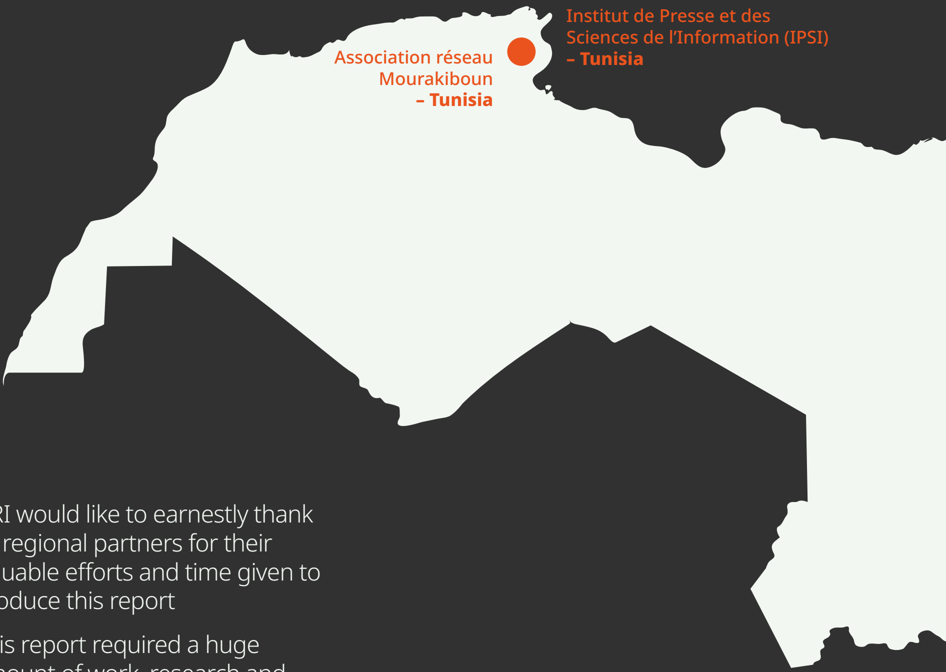
Jordan
Open Source
Association

Supported by

german
cooperation
DEUTSCHE ZUSAMMENARBEIT

September 2023

Acknowledgments


Association réseau Mourakiboun – Tunisia
Institut de Presse et des Sciences de l’Information (IPSI) – Tunisia

DRI would like to earnestly thank its regional partners for their valuable efforts and time given to produce this report

This report required a huge amount of work, research and dedication, and would not have been possible if we did not have the support of many individuals and organisations.

We would like, therefore, to extend our sincere gratitude to all of them.

Index

Maharat Foundation
– Lebanon
Jordan Open Source Association (JOSA)
– Jordan
Al Hayat Center (RASED) – Jordan

Executive Summary

This is the third of four regional social media monitoring reports to be produced as part of the "Words Matter" project, focusing on countering disinformation and hate speech in the MENA (Middle East and North Africa) region. The project aims to strengthen the safeguarding of democratic processes and the resilience of societies in the region to online disinformation and hate speech. It builds on the assumption that civil society actors, including the media, are essential to monitoring, understanding, and raising awareness of what debates and discourses are occurring online. This report is the result of the contributions of DRI's partners in Lebanon, Jordan, and Tunisia, who provided data-driven content related to their specific contexts from November 2022 to April 2023.

Words Matter's partner, Lab Track, presents a comprehensive report investigating the impact of disinformation and online gender-based violence on political discourse in the MENA region. That report focuses on the recent legislative elections in Tunisia and the social and legal context of gendered-hate speech in Jordan. Additionally, it showcases the awareness-raising efforts in Lebanon about the impact of disinformation and its influence on local elections, and includes an interview on the generation of political disinformation in Tunisia. The findings highlight the urgent need for collective action to address the growing challenges posed by disinformation and online harassment, in order to safeguard the integrity of democratic processes and promote gender equality.

The MENA region has been grappling with political instability, which has cast a concerning shadow over social media platforms and their content. The lack of independent media, comprehensive fact-checking mechanisms, and online literacy initiatives in the region contributes to the proliferation of misinformation. This uncontrolled spread of disinformation fosters an environment where extreme viewpoints dominate public discourse, hindering constructive dialogue and creating a polarising atmosphere. The lack of trust in traditional media

only exacerbates the problem, making it challenging for citizens to access reliable information and engage in informed discussions.

Online content moderation poses a significant challenge for social media platforms, especially given the diversity of languages and dialects spoken in the MENA region. Relying on automated content moderation and machine translation often leads to misinterpretations and inconsistent enforcement of content policies, leading either to harm or excessive censorship. This situation provides fertile ground for political actors to manipulate public opinion through deceptive tactics and visual content, such as memes and videos.

The report identifies three significant and persistent regional trends that impact the MENA region's political landscape.

Persistent Patterns of Disinformation in Election Campaigns:

In Tunisia, disinformation plays a pervasive role in undermining political candidates and questioning the integrity of the electoral system. False information is disseminated to create uncertainty and mistrust, aiming to influence public opinion, erode trust in the electoral process, and foster disillusionment among voters. On social media platforms, there is widespread use of fake accounts, masquerading as those belonging to supporters of specific candidates, while spreading disinformation against opponents. Moreover, popular Tunisian actors and television series are impersonated to mislead followers and manipulate public opinion.

In Lebanon, an unbalanced and distorted informational environment creates similar challenges in the electoral context. Disinformation spreads uncontrollably, impacting the electoral process and influencing public opinion. Here, again, the lack of independent media and comprehensive fact-checking mechanisms contribute to this unsafe environment, affecting 49.5 per cent of political discourse, primarily based on emotional political propaganda, rather than on factual information or policy debates.

The Strategic Use of Online Gender-Based Violence:

In Tunisia, the report reveals a concerning increase in online gender-based violence during electoral campaigns. Women candidates are disproportionately targeted with disinformation and hate speech, accounting for 45 per cent of such content directly addressing candidates, despite their representing only 11.5 per cent of the overall candidate pool. False narratives, personal attacks, and derogatory language are used to undermine the credibility and trustworthiness of women candidates, discouraging women from participating in politics and activism.

In Jordan, our partner, JOSA, observed online gender-based violence targeting women activists during the training phase of "Nuha", which is an AI model for detecting Online Gender Based Violence (OGBV) towards women in Jordan. Comments on social media platforms directed at women activists seek to intimidate and discourage them from participating in politics and activism. Misogynistic irony and sarcasm are employed to bypass content moderation. Cases documented by JOSA indicate that women politicians and activists, including feminist activists, are more subjected to hate speech online.

Hate speech against women in politics in **Jordan** includes demeaning comments, accusations of importing and implementing Western agendas, and the use of negative religious discourse and gender-based stereotypes. The report also includes a case study by our partner, Al Hayat Center – RASED, in Jordan, highlighting gender-based hate speech against Minister of Social Development Wafaa Bani Mustafa. The study found that 33.4 per cent of the comments on Facebook related to Bani Mustafa's statements contained gendered hate speech. These comments include negative religious discourse, gender-related stereotypes, and gender-based violence.

Coordinated Online Disinformation Campaigns During Major Political Events:

In Tunisia, the report reveals that networks of Facebook pages strategically coordinate their content and activities to amplify specific political positions during legislative elections. These coordinated efforts aim to influence the behaviour of Tunisian citizens and sway election outcomes in these actors' favour. Simultaneous live video broadcasts further highlight the coordinated nature of their activities. Facebook's algorithm favours live videos, leading to higher organic reach and engagement, making it challenging to detect hate content and disinformation through content-moderation filters. The report states that these same kinds of coordinated live video tactics have been observed during significant political events in the Tunisia in the past. Prior instances of coordinated live videos for misinformation and manipulation purposes have also been documented by the Words Matter network. Further research is needed to determine whether the same accounts and pages stand behind these networks.

To combat the challenges posed by disinformation and online gender-based violence, this report makes several key recommendations for different stakeholders:

Recommendations

Social Media Companies:

Social media platforms must invest more resources in the moderation of Arabic-language content. The prevalence of networks of fake accounts and morphing pages during elections necessitates dedicated country election task forces. Collaborating with third-party fact-checking initiatives can help identify patterns of coordinated sharing activities and enforce campaign standards. Platforms should work closely with official bodies in their countries to verify candidates' accounts during elections.

Governments:

Governments should invest in increasing social media literacy, to close the knowledge gap for different communities engaging and interacting online. Initiatives by civil society and national campaigns, supported by the government, can foster resilience to and awareness of disinformation. Social media education modules should be introduced in schools, to equip pupils with critical thinking skills and media literacy.

Civil Society:

Local civil society organisations (CSOs), media outlets, and researchers should connect and collaborate to establish a system of checks and balances for national authorities and tech platforms. Monitoring online public discourse during key events and providing evidence in findings can drive policy change. Ensuring researchers' mental health is vital, given the repeated exposure to harmful content. Support and resources for self-care and mental health should be made available to researchers monitoring disinformation and hate speech.

Researchers:

Further research on gendered hate speech in the region is crucial to informing effective policies in prevention, resource provision, enforcement of national regulations against harassment, improved detection by tech companies of hateful content in various Arabic dialects, and better enforcement of content-moderation guidelines.

A number of difficulties affected the research process at the regional level, including:

Limitations of research

The Military Conflict in Sudan:

The safety of the Sudan partner’s team was compromised due to intense clashes, hindering the implementation of their social media monitoring work.

The Internet Shutdown in Sudan:

An internet shutdown during the crisis in Sudan affected connectivity for millions of users, including our partner, limiting access to crucial information during a time when internet access is vital.

Changes in the Legal Framework:

Decrees and laws impacting the legal framework related to disinformation and hate speech in countries in the region raised concerns about potential repercussions for the publication of reports about disinformation during elections on social media platforms.

Restricted Access to Information:

Difficulty accessing essential information, such as the ages of candidates in Tunisia, posed challenges in filtering data for analysis. The inaccessibility or removal of social media content affected data analysis in both Tunisia and Jordan.

Changes in Social Media Platforms' Policies:

Changes in social media platforms' policies, such as access to Twitter's application programming interface (API), had negative technical implications for data retrieval and analysis.

Lack of Research on Online Gender-Based Violence in Jordan:

The lack of prior research in Jordan on online gender-based violence at a national level posed another significant challenge for Words Matter’s partner during the project.

Nonetheless, this report provides a comprehensive overview of the impact of disinformation and online gender-based violence on political discourse in the MENA region.
The trends and tactics identified highlight the urgent need to address these challenges collectively. By implementing the recommended strategies and overcoming research limitations, stakeholders can work together to create a healthier online environment, protect democratic processes, and promote gender equality. A collaborative effort is essential to safeguarding the region's democratic values and ensuring that social media platforms serve as platforms for informed discussions, a diversity of opinions, and meaningful civic engagement.

Introduction

In the first part of the report, Words Matter's partner, Lab Track, investigates the presence in online space of **Tunisian** legislative elections that took place in December 2022/January 2023, following the referendum in favour of the new Constitution in July 2022.

The second part of the report showcases the work of our partners in Jordan, the "Jordan Open-Source Association" (JOSA) and Al Hayat Center – RASED. The report discusses the legal and social context of gendered-hate speech in **Jordan**, as well as the results of research on patterns of gender-based hate speech. JOSA presents the results and insights gained from building the machine-learning module "Nuha", designed to detect online violence against women in Arabic in the Jordanian context. Meanwhile, Al Hayat Center – RASED analyses comments made in Facebook posts related to statements made by Wafaa Bani Mustafa, the Minister of Social Development in Jordan, in July 2022. The study sheds light on the extent to which gender-based hate speech is used to discredit and undermine women in leadership positions in Jordan and provides recommendations for addressing this issue.

The spotlight section of this report features the efforts of Words Matter partner in **Lebanon**, Maharat, in engaging in an inclusive literacy awareness campaign to pre-bunk false information before the local elections to be held in 2024. It also includes an interview with the Institut de Presse et des Sciences de l'Information (IPSI) about the generation of political disinformation in Tunisia.

Regional Context, Trends, and Findings

1. Regional context

Political instability in the MENA region casts a concerning shadow over social media platforms, making them a distorted and unbalanced environment. As indicated in the first[1] and second[2] social media monitoring reports produced by the Words Matter Network, in addition to certain cases highlighted in this report, engaging in political discussions on social media is often risky, due to the uncontrolled spread of disinformation, amplified by a lack of independent media, comprehensive fact-checking, and online literacy initiatives. This situation gives rise to extreme viewpoints dominating the discourse, hindering constructive dialogue, and creating a polarising atmosphere. The lack of trust in traditional media further contributes to the proliferation of misinformation, making it challenging to access reliable information and leading to uninformed discussions.

Words Matter has observed many instances of political disinformation on social media platforms in the MENA region, with some political actors manipulating public opinion through deceptive tactics. This trend extends beyond electoral contexts and permeates many online political discussions, fostering an environment where harmful narratives thrive, impacting the reputation of individuals and political parties. The uncontrolled spread of these narratives leads to harmful behaviors, including the propagation of hate speech by followers of these political actors.

Content moderation in Arabic poses a significant problem for social media platforms, particularly Meta (Facebook, Instagram), given the diversity of dialects spoken in the MENA region. Reliance on automated content moderation and machine translation often results in misinterpretations and inconsistent enforcement of content policies, leading

[1] DRI, "Online Public Discourse in MENA: Disinformation and Hate Speech During the 2022 Lebanese and Jordanian Elections", 28 September 2022.

[2] DRI, "Online Public Discourse in MENA: Regional Trends and Local Narratives", 21 February 2023.

either to harmful discourse or excessive censorship.

There has been a noteworthy shift in behavior in disinformation dissemination. Rather than relying solely on automated bots, political actors now depend heavily on visual content, such as memes, videos, and livestreams. This shift allows harmful narratives to spread organically on a larger scale, posing challenges for social media platforms in effectively monitoring and addressing deceptive campaigns. The impact of these campaigns is not limited to ordinary discussions, as events such as the legislative elections in Tunisia can also be influenced by social media campaigns, potentially affecting voters' perceptions, and compromising the integrity of the electoral process. Addressing these issues becomes crucial to maintaining a fair and transparent democratic system in the region.

In the last couple of months, the region has witnessed different challenges of economic and social natures. In Sudan, by mid-April, intense clashes between Sudan's Military Forces (SAF) and the country's main paramilitary force, Rapid Support Forces (RSF), had killed hundreds of people and sent hundreds of thousands more fleeing for safety, as the burgeoning civil war threatened to destabilise the wider region.

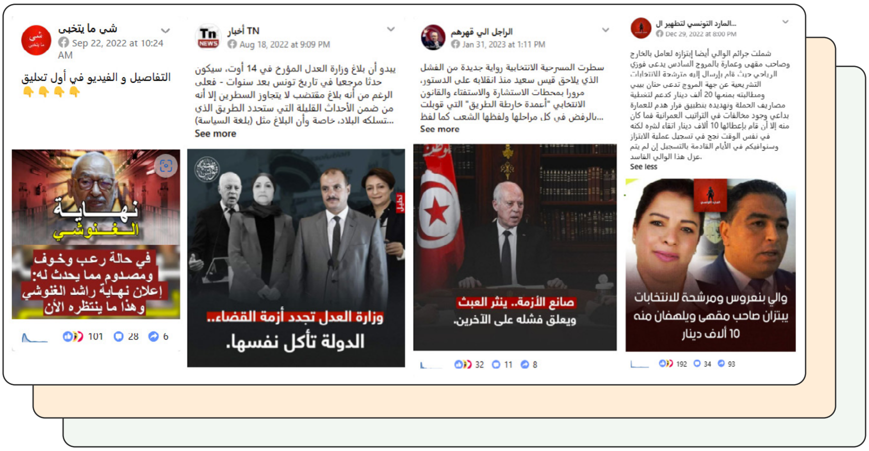
This extreme change in the context of Sudan has prevented the participation of Words Matter's Sudanese partner in this report. Both military forces in Sudan have been using social media tools and techniques to push their own narratives inside of the country, as well as to target the international community. Social media and "Open-Source Intelligence" researchers investigated the main tactics used in online manipulation during the crisis in Sudan.

2. Regional Trends and Tactics

First regional trend: Persistent disinformation patterns cross countries to manipulate elections.

The first regional trend in **Tunisia** revolves around the pervasive use of disinformation for political manipulation, targeting not only candidates, but also electoral systems themselves. Such content is rife with propaganda, conspiracy theories, and political manipulation, creating an atmosphere of uncertainty and mistrust. In this context, false information is disseminated to undermine the credibility of candidates, casting doubt on their qualifications and intentions. Additionally, disinformation is used to challenge the electoral system, questioning its fairness and integrity. By sowing the seeds of doubt and confusion, those propagating disinformation aim to influence public opinion, erode trust in the electoral process, and create a sense of disillusionment among voters.

The tactics employed in this trend include **the use of fake accounts** on social media platforms. In **Tunisia**, such accounts have falsely claimed to be those of supporters of specific candidates but were, in fact, spreading disinformation against other candidates. Moreover, some of these fake accounts have impersonated well-known Tunisian actors, animators, and popular Tunisian television series. By using the popularity and influence associated with these individuals and programmes, these fake pages aimed to mislead their followers and manipulate public opinion.



شي ما يتخبى
Sep 22, 2022 at 10:24 AM

التفاصيل و الفيديو في أول تعليق
👇👇👇👇

101 28 6

أخبار TN
Aug 18, 2022 at 9:09 PM

يبدو أن بلاغ وزارة العدل المؤرخ في 14 أوت، سيكون حدثا مرجعيا في تاريخ تونس بعد سنوات - فعلى الرغم من أنه بلاغ مقتضب لا يتجاوز السطرين إلا أنه من ضمن الأحداث القليلة التي ستحدد الطريق الذي تسلكه البلاد، خاصة وأن البلاغ مثل (بلغة السياسة...
See more

الراجل الي قهرهم
Jan 31, 2023 at 1:11 PM

سطرت المسرحية الانتخابية رواية جديدة من الفشل الذي يلاحق قيس سعيد منذ انقلابه على الدستور، مرورا بمحطات الاستشارة والاستفتاء والقانون الانتخابي "أعمدة خارطة الطريق" التي قوبلت بالرفض في كل مراحلها ولفظها الشعب كما لفظ...
See more

32 11 8

المارد التونسي لتطهير ال...
Dec 29, 2022 at 8:00 PM

شملت جرائم الوالي أيضا إبتزازه لعامل بالخارج وصاحب مقهى وعمارة بالمروج السادس يدعى فوزي الرياحي حيث قام بإرسال إليه مترشحة الانتخابات التشريعية عن جهة المروج تدعى حنان بيبي ومطالبته بمنعها 20 ألف دينار كدعم لتغطية مصاريف الحملة وتهديده بتطبيق قرار هدم للعمارة بداعي وجود مخالفات في التراتيب العمرانية فما كان منه إلا أن قام بإعطائها 10 آلاف دينار اتقاء لشره لكنه في نفس الوقت نجح في تسجيل عملية الابتزاز وسنوافيكم في الأيام القادمة بالتسجيل إن لم يتم عزل هذا الوالي الفاسد.
See less

192 34 93

نهاية الغنوشي
في حالة رعب وخوف ومصدوم مما يحدث له: إعلان نهاية راشد الغنوشي وهذا ما ينتظره الآن
تحليل
وزارة العدل تجدد أزمة القضاء..
الدولة تأكل نفسها.
صانع الأزمة.. ينثر العبث
ويعلق فشله على الآخرين.
والي بنعروس ومرشحة للانتخابات يبتزان صاحب مقهى ويلهفان منه 10 ألاف دينار

Figure 1: An example of Tunisian disinformation content, involving a manipulative tactic that plays on voters' opinions.

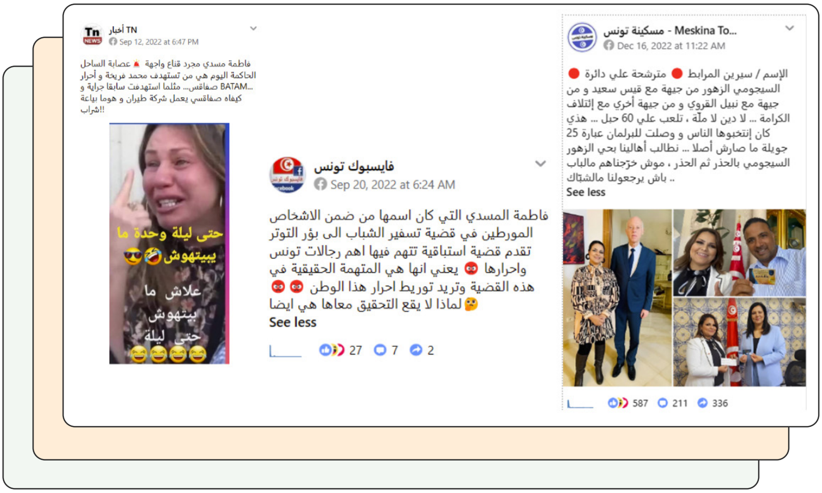
These examples display false news about candidates and conspiracy theories. The first screen shot (right to left) particular screenshot claims that “ "Hanen Bibi", a candidate for parliamentary elections, was involved in blackmaling of a coffee shop owner with Ben Arous governor to take over the shop”. Another screenshot presents President Kais Saied with a caption as a “troublemaker who spreads chaos and pins his failure on others” with above text describing the elections as a “charade”. . The third screenshot claims that the Minister of Justice is aggravating problems within the judicial system. In Lebanon, the trend of disinformation during parliamentary elections is characterised by an unbalanced and distorted informational environment, creating an unsafe space for political and public discussions. This unsafe environment is exacerbated by the lack of independent media, comprehensive fact-checking mechanisms, and digital information and media literacy initiatives. Consequently, disinformation spreads uncontrollably, aiming to influence public opinion and to impact the electoral process. These results are similar to the findings presented by Words Matter partner Maharat in the second regional report, during the parliamentary elections in 2022, which showed that 49.5 percent of the political discourse online was based on emotional political propaganda, instead of being based on factual information or political arguments to support political positions or to counter or confront political opponents. Hate speech and violence against women in political and public life formed a large part of the public discourse.

The tactic of creating fake accounts around the legislative elections in both Lebanon and Tunisia was respectively documented by the Words Matter Network in DRI's second and third regional report.

Second regional trend: The strategic use of online gender-based violence.

In Tunisia, the team observed a worrying increase in online gender-based violence during electoral campaigns. The data reveals that women candidates are disproportionately the targets of disinformation and hate speech, accounting for 45 per cent of all content directly addressing candidates, even though they make up only 11.5 per cent of the overall candidate pool.

Examples of online gender-based violence include posts that accuse women candidates of involvement in illegal activities or blame them for various issues. False narratives, personal attacks, and derogatory language are used to undermine the credibility and trustworthiness of women candidates. For instance, once candidate, **Fatma Mseddi**, has been accused of sending young people to Syria, while another, **Syrine Mrabet**, was attacked for her associations with influential individuals.



TN أخبار
Sep 12, 2022 at 6:47 PM

فاطمة مسدي مجرد قناع واجهة 🚨 عصابة الساحل
الحاكمة اليوم هي من تستهدف محمد فريخة و أحرار
صفاقس... مثلما استهدفت سابقا جراية و BATAM...
كيفاه صفاقسي يعمل شركة طيران و هوما بياعة
شراب!!

فايسبوك تونس
Sep 20, 2022 at 6:24 AM

فاطمة المسدي التي كان اسمها من ضمن الاشخاص
المورطين في قضية تسفير الشباب الى بؤر التوتر
تقدم قضية استباقية تتهم فيها اهم رجالات تونس
واحرارها 😡 يعني انها هي المتهمة الحقيقية في
هذه القضية وتريد توريط احرار هذا الوطن 😡 😡
🤔لماذا لا يقع التحقيق معاها هي ايضا
See less

27 7 2

مسكينة تونس - Meskina To...
Dec 16, 2022 at 11:22 AM

الإسم / سيرين المرابط 🔴 مترشحة علي دائرة 🔴
السيجومي الزهور من جيهة مع قيس سعيد و من
جيهة مع نبيل القروي و من جيهة أخري مع إئتلاف
الكرامة ... لا دين لا ملّة ، تلعب علي 60 حبل ... هذي
كان إنتخبوها الناس و وصلت للبرلمان عبارة 25
جويلة ما صارش أصلا ... نطالب أهالينا بحي الزهور
السيجومي بالحذر ثم الحذر ، موش خرّجناهم مالباب
باش يرجعولنا مالشبّاك ..
See less

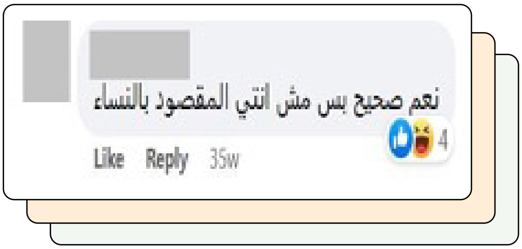
587 211 336

Figure 2: Screenshots capturing posts containing hate speech/ disinformation/ physical appearance mockery.

In Jordan, Words Matter partner **JOSA** observed **online gender-based violence targeting women activists** during the language-training phase of the "Nuha "model. The model picked up death threats, insults, bullying, and the propagation of stereotypes targeting women activists. Such comments directed at women activists aim not only to undermine their credibility, but also resort to derogatory language, sexualising their appearance, and making personal attacks. The malicious intent behind these comments seeks to intimidate and discourage women from participating in politics and activism, with the share of online gender-based violence at 32 per cent of all the comments on posts by women activists.

The trend of online gender-based violence in Jordan is further exacerbated by the use of **misogynistic irony and sarcasm**, **often used to escape content moderation.**

Al Hayat Center – RASED documented misogynist comments laced with sarcasm often label women candidates as "not women" as they engage in politics, which is seen as a man’s job, implying that they are not suitable for political positions.



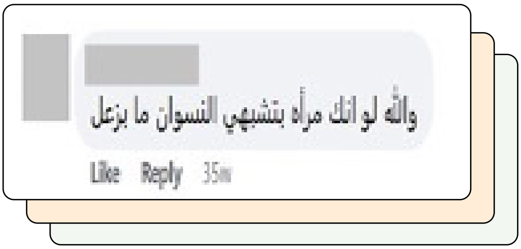
[redacted] [redacted]

نعم صحيح بس مش انتي المقصود بالنساء

4

Like Reply 35w

“Yes, that's correct, but you are not included in this because this is only directed towards women.”

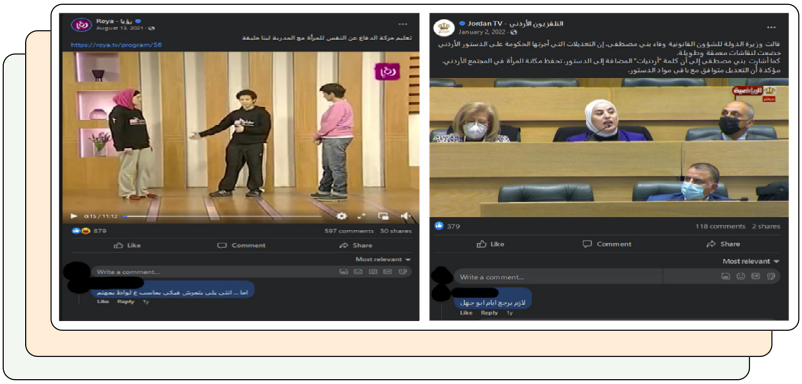


[redacted] [redacted]

والله لو انك مرأه بتشبهي النسوان ما بزعل

Like Reply 35w

“If you were a woman that actually looked like women, I would not be as angry.”



Roya - رؤيا

August 13, 2021

تعليم حركة الدفاع عن النفس للمرأة مع المدربة لينا خليفة

https://roya.tv/program/56

0:15 / 11:12

879 597 comments 30 shares

Like Comment Share

Most relevant

[redacted] Write a comment...

[redacted]

اما .. انتي يلي بتحرش فيكي بحاسب ع لواط بجهنم

Like Reply 1y

Jordan TV - التلفزيون الأردني

January 2, 2022

قالت وزيرة الدولة للشؤون القانونية وفاء بني مصطفى، إن التعديلات التي أجرتها الحكومة على الدستور الأردني خضعت لنقاشات معمقة وطويلة.
كما أشارت بني مصطفى إلى أن كلمة "أردنيات" المضافة إلى الدستور، تحفظ مكانة المرأة في المجتمع الأردني، مؤكدة أن التعديل متوافق مع باقي مواد الدستور.

379 118 comments 2 shares

Like Comment Share

Most relevant

[redacted] Write a comment...

[redacted]

لازم يرجع ايام ابو جهل

Like Reply 1y

Figure 3: An example of insulting and bullying comments Stereotyping the trainer Lina Khalifa (left), and of comments with death threats targeting minister Wafaa Bani Mustafa (right)

The dataset JOSA analysed revealed that hate speech against women in politics in Jordan takes one of two forms: either women are told to “return to the kitchen, where they belong” – a literal phrase commonly used in Jordan – or they are accused of importing and implementing Western agendas that aim to destroy the Jordanian community.

This finding is similar to that in the social media research conducted by DRI in Libya in 2022 – "Firmer Ground for Advancing Women's Participation in Libya: Social Media Report" – which found "get back to the kitchen" to be one of the most repeated key phrases found in the hate speech used against women politicians on social media.

The case study conducted by Words Matter partner Al-Hayat Center – RASED in Jordan on gender-based hate speech against Social Development Minister Wafaa Bani Mustafa provides valuable insights into the trend of online gender-based violence in Jordan. The study found that 33.4 per cent of the comments on Facebook related to Bani Mustafa's statements contained gendered hate speech.

The interactions featured disapproving questions about gender equality, demeaning comments, accusations of foreign influence, hate speech targeting the minister's appearance, and concerns about the impact of promoting gender equality. The comments included negative religious discourse, gender-related stereotypes, and gender-based violence, such as sexual harassment and misogyny. Addressing this trend requires raising awareness, strong content moderation, and creating support systems to protect women from online violence and encourage their meaningful participation in public life without fear of harassment or discrimination.

This trend reinforces the findings of the second regional report, where Words Matter Network partners documented gendered hate speech during elections, and shoed that comment sections are where most of the hate speech targeting women's public statements appears.

Third regional trend: Coordinated online disinformation campaigns during major national political events.

In Tunisia, our research revealed networks of Facebook pages strategically coordinating their content and activities to amplify specific political positions, such as advocating for the "yes" or "no" vote in the referendum on the new Constitution proposed by President Kais Saied, or for voters to refrain from going to the polls. The ultimate objective of these coordinated networks was to influence the political behavior of Tunisian citizens and sway the election's outcome in favor of their preferred candidatesd identical posts with the same format, writing style, and even video presentations, indicating a deliberate effort to synchronise the dissemination of this content. The discovery of simultaneous live video broadcasts further emphasised the coordinated nature of their activities.

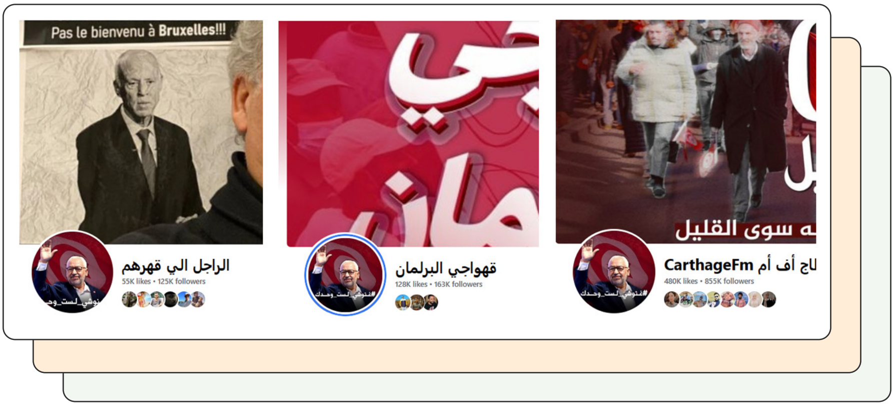

Pas le bienvenu à Bruxelles!!!
الراجل الي قهرهم
55K likes • 125K followers
قهواجي البرلمان
128K likes • 163K followers
#غنوشي_لست_وحدك
ه سوى القليل
CarthageFm لاج أف أم
480K likes • 855K followers
#غنوشي_لست_وحدك

Figure 4: Examples of different pages sharing the same profile picture and involved in coordinated disinformation campaigns.

During the Tunisian legislative elections, our researchers observed a coordinated live video tactic. Facebook's algorithm favours live videos, notifying page and group followers in advance, leading to higher organic reach and increased engagement compared to pre-recorded content. The simultaneous broadcast of live videos on multiple pages made it challenging to detect disinformation they contained through platform content moderation filters. This trend allowed certain actors to leverage live videos to spread disinformation and manipulate public opinion during the elections, highlighting the importance of effectively monitoring and addressing the impact of live video content on social media platforms.

The Words Matter network has thus documented the use of the same tactic of posting coordinated live videos on multiple pages during two significant political events in the same country. This practice during the legislative elections was also seen in the campaign for the referendum in July 2022, covered in the second regional report. It is important to note, however, that these aren't isolated events. Prior to this report, the network had already identified instances of coordinated live videos being used in other contexts for misinformation and manipulation purposes. Further research into whether the same accounts and pages are behind this network is further needed.

3. Regional Recommendations for social media companies

Invest greater resources in the moderation of Arabic-language content. The Words Matter partners, once again, collected evidence about the use of networks of fake accounts and morphing pages around the time of elections, with a limited ability to identify them and delayed responses from social media platforms in acting against them. We call on social media platforms to assign dedicated election task forces to each country, to collaborate with third-party fact-checking initiatives in focusing investigations on patterns of coordinated sharing activities by pages and groups, and to establish campaign standards with official country bodies to verify candidates' campaign-related accounts.

For governments:

Invest in increasing social media literacy. Closing the knowledge gap among different communities engaging and interacting online is crucial to inoculating citizens against disinformation. This could take the form of civil society initiatives and national campaigns, supported by governments, to foster resilience to and awareness of disinformation and other forms of online manipulation. Social media education modules should be introduced in school curricula.

For civil society:

Local CSOs, media outlets, and researchers should connect, collaborate, and establish a system of checks and balances for national authorities and tech platforms. National level authorities (parliaments, law enforcement, media observation agencies, electoral monitoring boards, etc.) and tech companies often show either a lack of familiarity with the issues at stake or a lack of interest in working towards ensuring a healthier online environment. Local media and CSOs should join forces to systematically monitor online public discourse leading up to and during key events in the region, and to produce the evidence necessary to push for changes to policies and practices.

Put the mental health of researchers first. Words Matter partners noted that their social media monitors and researchers were negatively affected by the repeated exposure to harmful content. Sifting through disturbing material can induce trauma. CSOs, media organisations and human rights advocates that monitor online disinformation and hate speech should anticipate this and ensure that support and resources are available for their members, including training on self-care and mental health. Donors that fund social media monitoring work should understand this as a legitimate cost.

For researchers:

Conduct large-scale research on gendered hate speech in the region. There is a clear pattern of hate speech and harassment specifically targeting women in the region, differing in nature and intensity depending on the country context. This needs to be the focus of further research, to inform the development of effective policies in (i) prevention; (ii) provision of resources for targets; (iii) enforcement of national regulations against harassment; (iv) improved detection by tech companies of hateful content in various Arabic dialects; and (v) improved enforcement of platforms' own guidelines.

4. Research Limitations at the Regional Level

The military conflict in Sudan threatened the safety and of partner Sudia's team and prevented them from carrying out their work on social media.

The shutdown of the internet in Sudan impacted the connectivity of millions of users in Sudan, including our partner, at a time when internet access is crucial to saving lives.

Decrees and laws have changed the legal framework related to disinformation and hate speech in countries in the region, such as Decree 54 in Tunisia and the amended “Electronic Crimes Law” that is currently being discussed in Jordan, which includes broad clauses that would grant the government the ability to charge individuals with crimes that could lead to jail for several years and fines of thousands of dollars. This creates the fear that publishing reports about disinformation related to elections on social media platforms during elections might lead authorities to ban a social media platform, or to shut down the internet during elections.

Restricted access to information was a significant limitation.

- In **Tunisia**, Words Matter’s partner had intended to include the ages of candidates as an important criterion in data analysis, but it was extremely difficult to obtain the data on time from the Independent High Authority for Elections (ISIE), making it so the data could not be filtered according to this criterion.
- In **Tunisia** and **Jordan**, a significant proportion of posts and comments were either non-accessible or had been removed. The inaccessibility of such content can pose challenges and hamper the data analysis process. Words Matter’s partners reported difficulties accessing the Twitter API, following Elon Musk’s takeover of the company and decision to revoke access to its free API. JOSA had to resort to an alternative tool, called “Export Comments”. The situation with Twitter's policy changes not only affected Words Matter partners, but also had broader technical implications for other third-party applications that relied on Twitter's API for data retrieval and analysis.
- In **Jordan**, the lack of prior research on online gender-based violence at a national level posed another significant challenge for Words Matters’ partner during the project. The absence of adequate academic research in Jordan on the intersections of gender and artificial intelligence made it incumbent on the team to start almost from scratch, instead of building on existing research findings, which has ultimately contributed to the difficulty in defining and operationalising key hate speech-related concepts.

Country Case Studies

1. 2022-2023 Tunisian Legislative Elections: Analysing Online Public Discourse on Social Media

Introduction

This report aims to investigate the disinformation and hate speech trends in the Tunisian online space during the 2022-2023 legislative elections, to inform the debate between civil society stakeholders on actors involved, causes, and drivers of these trends, and to produce recommendations to countering these, while paving the way for further research by relevant stakeholders.

The analysis focuses on harmful behaviors observed on Facebook during the electoral period, including inauthentic networks, coordinated campaigns for spreading disinformation, targeting candidates, and morphing pages.

The monitoring covered over 477 Facebook pages during both rounds of the elections. The period for the first round was from 25 November to 17 December 2022, and that for the second round from 16 January to 29 January 2023.

The data collection phase extended from 1 August 2022 to 27 February 2023, starting one month before the first round of voting and ending one month after the second.

The online campaigning for the elections comprised not only calls and campaigns to vote for specific candidates, but also calls to boycott the vote.

Context

This report focuses on the first and second rounds of the Tunisian legislative elections in December 2022 and January 2023, respectively. The exceptional measures initiated by President Kais Saied in July 2021 culminated in the dismissal of the democratically formed government, the dissolution of the previously elected parliament, and the introduction of a new constitution, marking a major regression from the country's democratic gains made after the 2011 revolution and the 2014 Constitution.

The new electoral law,[3] introduced by presidential decree 55 in September 2022, brought about significant changes to the electoral system. It shifted Tunisia from a party-based electoral system to one based on individual candidatures for the Assembly of Representatives of the People (ARP).

The rise of political polarisation in Tunisia poses a threat to freedom of expression and democratic development. Article 24 of Decree-Law 54[4] states that anyone who knowingly uses information and communication systems to spread false news, data, or documents, with the aim of harming public safety, national defense, or infringing on the rights of others, will be punished. The law is punishable by imprisonment for five years and a fine of 50,000 Tunisian Dinars, or 16,100 U.S. dollars.

The problem with the law lies in the subjective power that is given to the authorities to determine what constitutes "false news" or information. This subjectivity opens the door to selective enforcement, where authorities may target individuals or groups based on their political views. When applied in a biased manner, the law could be used as a tool to suppress dissenting opinions and stifle political opposition, further exacerbating political polarisation within the country.

Furthermore, the fear of legal repercussions for expressing one's opinions can lead individuals to self-censor and refrain from engaging in open discussions or expressing alternative viewpoints. The potential consequences of imprisonment and hefty fines create a chilling effect on free speech, silencing individuals and discouraging them from participating in public discourse. This silencing effect contributes to the formation of isolated ideological bubbles and echo chambers, where people primarily interact with those who share similar political beliefs. As a result, political polarisation becomes more pronounced, as individuals have limited exposure to diverse perspectives and are less inclined to engage in constructive dialogue with those who hold opposing views. The National Syndicate of Tunisian Journalists (SNJT) has called on the president of the Republic to withdraw the decree, so as to avoid any infringement on the freedom of expression, stating that the law is part of legal measures being imposed to restrict freedom of speech.

Methodology and data approaches

The methodology employed in this research was data-driven, utilising a systematic approach to collecting, analysing, and classifying social media data from Facebook during the Tunisian legislative elections.

[3] Sarah Yerkes & Mohammad Al-Mailam, "Tunisia's New Electoral Law Is Another Blow to Its Democratic Progress", Carnegie Endowment for International Peace, 11 October 2022.

[4] DCAF Tunisie, "Décret-loi n° 54-2022 du 13 septembre 2022, relatif à la lutte contre les infractions se rapportant aux systèmes d'information et de communication", 13 September 2022.

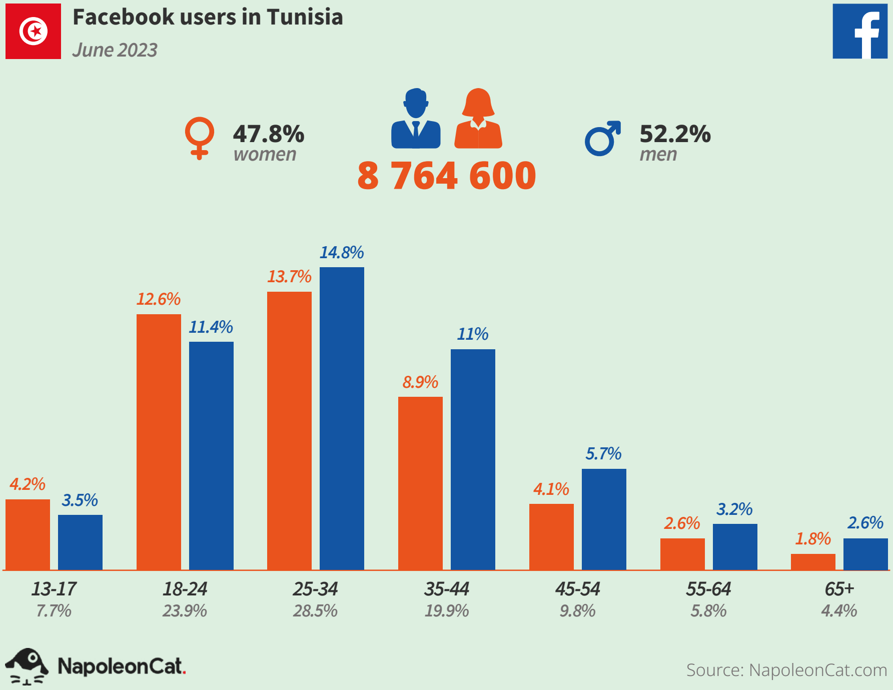
The sample selection process involved careful criteria for choosing the social media platform, which was based on four aspects that are mentioned in detail below. For data gathering, research used "CrowdTangle", a social media analytics tool that allows for the collection and analysis of valuable data from social media platforms to provide insights into posting activity and engagement metrics. The collected data was then annotated with the help of Label Studio, an open-source free machine learning (ML) tool to classify data into categories such as verified information, not verified information, hate speech, misinformation, disinformation, and sarcasm, allowing for a deeper understanding of the online discourse during the elections.

This data-driven methodology combines both qualitative and quantitative analysis, to ensure a comprehensive and evidence-based understanding of social media dynamics in relation to the electoral process in Tunisia.

Sample Selection

1. Selection of Social Media Platform:

Facebook was chosen as the primary social media platform for data collection and analysis, due to its widespread usage in Tunisia, where some 8.7 million Facebook accounts are registered, accounting for over 70 per cent of the population.



Source: NapoleonCat.com

Figure 5: Number of Facebook users in Tunisia – according to NapoleonCat.com

2. Facebook Pages Selection: A total of 477 Facebook pages were selected for monitoring during both rounds of the elections.

The selection criteria for pages considered various factors, including:

- Mainstream media outlets: Pages belonging to official media outlets to capture news coverage of the elections. These pages belong to established and widely recognised media organisations that are considered authoritative sources of news. They include television channels, radio stations, news websites, magazines, and media companies. These outlets typically adhere to professional journalistic standards and provide news coverage, analysis, and reporting on a wide range of topics, including politics, current events, and elections. Their primary focus is to provide objective and balanced news coverage, reaching a broad audience.
- Governmental pages: Pages associated with government institutions.
- Politics-/news-related pages: Pages dedicated specifically to political discussions, analysis, and news commentary. These may include pages run by journalists, political analysts, bloggers, or individuals interested in politics. Unlike mainstream media outlets, these pages might have a specific political leaning or focus on a particular aspect of politics or news coverage. They often offer commentary, opinions, and in-depth analysis of political events, policies, and election campaigns. These pages may attract a more niche audience interested in political discourse and analysis.
- CSO pages: The pages of CSOs working in the field of democracy and/or engaged in monitoring the electoral process.
- Political parties and candidates: Pages of political parties and individual candidates participating in the elections.
- Public pages with political content: Random pages that are not necessarily identified as political, but actively share political content.

PAGE CLASSIFICATION	NUMBER OF PAGES	PERCENTAGE
Mainstream media outlet pages	100	35.6%
Governmental pages	38	7.9%
Politics-/news-related pages	74	15.5%
CSO pages	12	2.5%
Political parties/candidates	53	11.11%
Public pages with political content	200	41.9%
Total	477	100%

This hand-picked selection of the accounts list aimed to encompass official sources from parties and official institutions, CSOs, and official media covering the elections, for a comprehensive understanding of online political discourse during the campaign.

3. Keyword Selection:

In addition to the pages, a list of keywords was included to target specific content related to the elections. This may include the names of candidates (قائمة المترشحين، مترشح...), election-related terms (for example تصويت، اقتراع، صناديق، حملة انتخابية), and other relevant keywords (قانون انتخابات جديد، قانون انتخابات قيس سعيد) to capture a comprehensive range of election-related discussions. (see: Annex 1, p.40)

4. Date Selection:

The data collection period spanned from August 1st, 2022 to February 27th, 2023. This period covers from one month before the beginning of the first round, allowing for the analysis of pre-election campaign discourse, and one month after the second round, to capture post-election discussions and any potential shifts in social media behavior.

Data Gathering

The data was gathered using the CrowdTangle tool, which allows for the collection of large amounts of data from various social media platforms, including Facebook. It provides insights into engagement metrics, post activity, and other relevant information.

From the 477 pages monitored, we collected 11,567 posts fitting the various categories mentioned above.

Data Annotation and Classification:

The collected data was subjected to annotation and categorised based on the type of speech that each example represented.

A team of annotators with a deep understanding of the Tunisian context was responsible for carrying out the task. To ensure thorough verification, the annotators investigated each text column by cross-referencing the available URLs and page names associated with the posts. By examining the provided URLs, the annotators could access the specific web pages or online sources where the posts originated. This allowed them to gather additional information, validate the content, and assess its credibility.

The following manual classification categories were utilised:

Main content verification skills were applied by a team of annotators throughout the timeline of the elections, applying the five pillars of verification practice (provenance, source, date, location, motivation) assessing 11,567 Facebook posts, dividing them into two main categories based on the below.

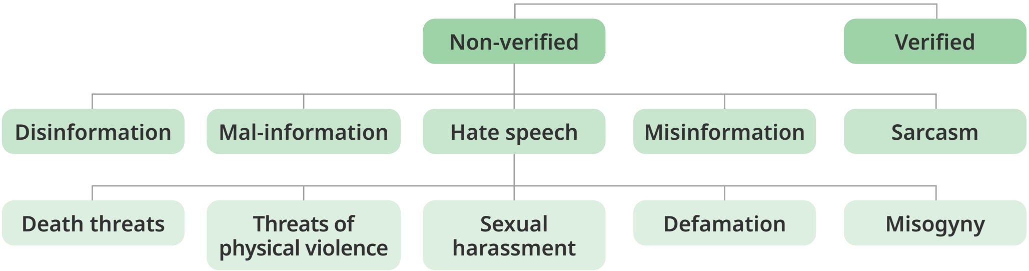


Figure 6: Data classification categories used in the manual annotation process

The definition of each category is presented below.

CATEGORY	DEFINITION
Verified	Information that has been verified to be true and accurate.
Non-verified	Claims that have not been verified or fact-checked.

Classification of non-verified content

Misinformation	False information that is spread, regardless of whether there was an intent to mislead.
Disinformation	False information that is spread with harmful intent and in an organised way.
Mal-information	Any kind of communication that involves leaks, doxing, or the release of private information; can be true or false.
Sarcasm	Communication that uses irony, mocking, or satirical remarks to convey a different meaning than a literal interpretation.
Hate speech	Any kind of communication in speech, writing, or behavior that attacks or uses pejorative or discriminatory language towards a person or group based on their identity (e.g., religion, ethnicity, gender).

Sub-classification of hate speech induced discourse

Death threats	Communications that include clear death threats.
Threats of physical violence	Communications that threaten physical violence.
Sexual harassment	Communications that contain expressions of misogyny and/or sexual harassment content.
Defamation	Communications that aim to defame someone's reputation.
Misogyny	Communications that focus on gender stereotypes and gender-based prejudice.

Another round of data annotation was conducted using Python notebooks, employing code-driven annotation. Python notebooks, such as Jupyter Notebooks, are interactive computing environments that allow for the execution of code, data analysis, and documentation in a single interface. These notebooks provide a convenient way to combine code snippets, visualisations, and explanatory text.

The code-driven annotation process involves using Python code to automatically annotate the data based on predefined rules, algorithms, or patterns. This approach eliminates the need for manual annotation, and enables efficient and consistent processing of the dataset.

During this annotation round, the focus was on specific aspects of the data. The code-driven annotation aimed to identify and extract relevant information related to these aspects, using programming logic and techniques:

- Timeline of elections: Notes and definitions related to different stages of the electoral process, including the pre-election period, the campaigning period, voting days (phase 1 and phase 2), and post-election results.

CATEGORY	DEFINITION
October 17	Opening of candidate submissions for legislative elections
October 24	Deadline for candidate submissions, at 6:00 PM
October 31	ISIE decides on candidacies
November 1	Publication of candidate lists at ISIE premises
November 21	Announcement of the final candidate list
November 25	Start of the electoral campaign in Tunisia and abroad
December 13	End of the electoral campaign abroad
December 14	Electoral silence abroad
December 15	End of the electoral campaign in Tunisia
December 16	Electoral silence in Tunisia
December 15-17	Elections abroad
December 17	Elections in Tunisia
December 20	Announcement of the preliminary results

CATEGORY	DEFINITION
January 19, 2023	Announcement of the final results
January 16	Start of the electoral campaign in Tunisia and abroad
January 27	End of the electoral campaign
January 28	Electoral silence in Tunisia
January 29	Elections in Tunisia
February 1	Announcement of the preliminary results
March 4	Announcement of the final results

- Geography: A dataset containing geographical information was utilised, including all of Tunisia's regions and constituencies as announced on the official ISIE website.

The code-based annotation process involves a natural language processing technique that scans all of the posts, analysing the "Text of post" column, "Link Text" if it is a link, or "Image Text" if the post contains an image. The code compares the text in these columns with the geographic dataset to identify any matches.

- Candidate Mentions: The data classification is based on a complete list of candidates. The process of classification involves determining whether a candidate is mentioned in the text columns related to each post. The goal is to identify posts that reference specific candidates, allowing for further analysis and insights into their visibility, popularity, or public perception.

The code utilises natural language processing (NLP) techniques to scan the text columns of each post, employing exact string matching and case-insensitive matching to identify mentions of candidate names, thus classifying posts as either mentioning a candidate or not.

- Gender of candidates: The gender of candidates was determined automatically, and they were classified as either women or men.

Based on the list of candidates provided by the ISIE, only 11.5 per cent of the candidates were women.

There were 122 women candidates in round 1 and 34 in round 2.

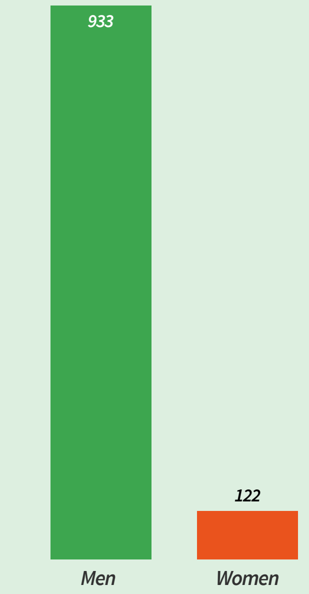


Figure 6: Comparing the number of men and women candidates in the elections (Source: ISIE)

- Type of content: The content was categorised automatically into different types, including text, link, video, live video, and image.

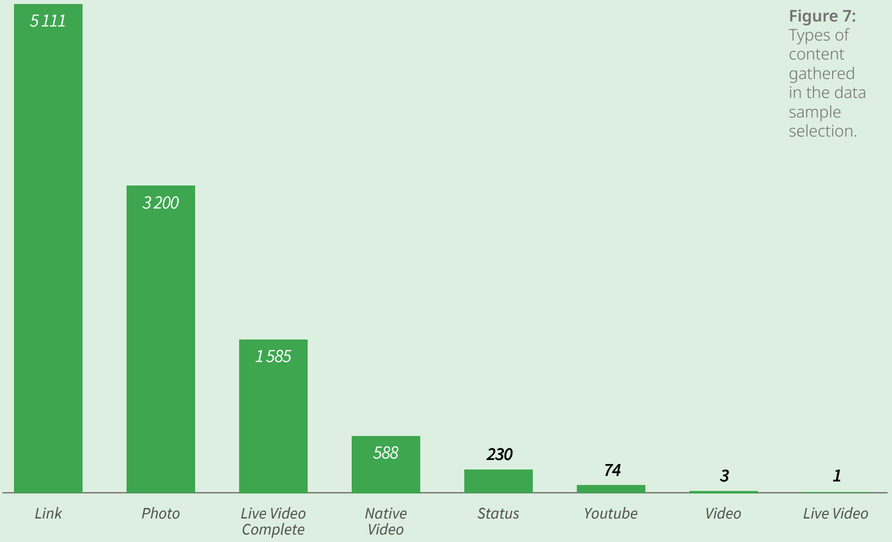


Figure 7: Types of content gathered in the data sample selection.

Data and Network Analysis:

During the data analysis of the content, the tools used were Python and Power BI.

By combining the features provided by these tools, the team were able to clean and process the data, and map the networks that were identified as coordinating the dissemination of similar content through visually compelling presentations.

- **Network identification of non-verified content methodology**

In the context of the data analysis, a network identification methodology was employed to identify networks of pages that shared similar posts. The objective was to explore potential collaborations or coordinated efforts among these pages in publishing non-verified content. The methodology focused on the following criteria:

1. Similar posts: Pages that shared more than three similar posts were considered for network identification. This criterion ensured that there was a significant overlap in the content being shared among the pages. By analysing these similarities, the team aimed to uncover connections and patterns in the dissemination of information.
2. Same-day publishing: The code specifically looked for instances where the shared posts were published on the same day. Moreover, the publishing times of these posts were observed to be very close to each other. This criterion allowed for the identification of pages that not only shared similar content, but also did so within a short time frame. This indicated potential coordination or collaboration in the timing of post publication.
3. Rule of exclusion of radio, news, and official pages: To focus on non-verified content and random pages, the methodology excluded radio, news, and official pages. By excluding these categories, it was ensured that the analysis primarily targeted pages that were not officially recognised media outlets or sources of verified information. This allowed for a deeper exploration of networks operating outside of established channels.

Research Limitations:

1. We intended to include the age of candidates as an important factor in filtering the results, particularly to identify young candidates. However, despite extensive research, we were unable to find information about the age of candidates on the ISIE website or on any radio or news sites, or even in forums for direct speech videos where candidates could express their views directly to the public. Currently, we only have information about candidates in the second round of the elections.
2. Content: During the process of analysis of the data, the team found many links that weren't accessible. (Three per cent of the data sample). This percentage may be due to various factors. Primarily, some posts were marked as such because their content had been deleted or removed,

rendering them inaccessible for verification purposes. Additionally, a portion of these posts originated from official news sources, radio stations, and pages associated with the ISIE. Given the trusted nature of these sources, the content was considered reliable and did not require independent verification. Furthermore, a subset of the posts contained neutral campaign messages, implying that their neutrality and lack of bias made additional verification unnecessary. Another aspect to consider was the possibility that certain pages changed their privacy settings from public to private, thereby restricting access to links and content within the posts for specific audiences. This change in status further hindered the ability to verify the affected posts.

Examples of non-accessible links on two different levels:

- Inaccessible links due to unavailability at the level of the page itself:

  This can happen if the page is no longer accessible or if its privacy settings have been changed, restricting access for certain audiences.

Page: **ما يفوتك شي**:

https://www.facebook.com/100084479785416/posts/144120631747262

https://www.facebook.com/100084479785416/posts/144429801716345

https://www.facebook.com/100084479785416/posts/144120631747262

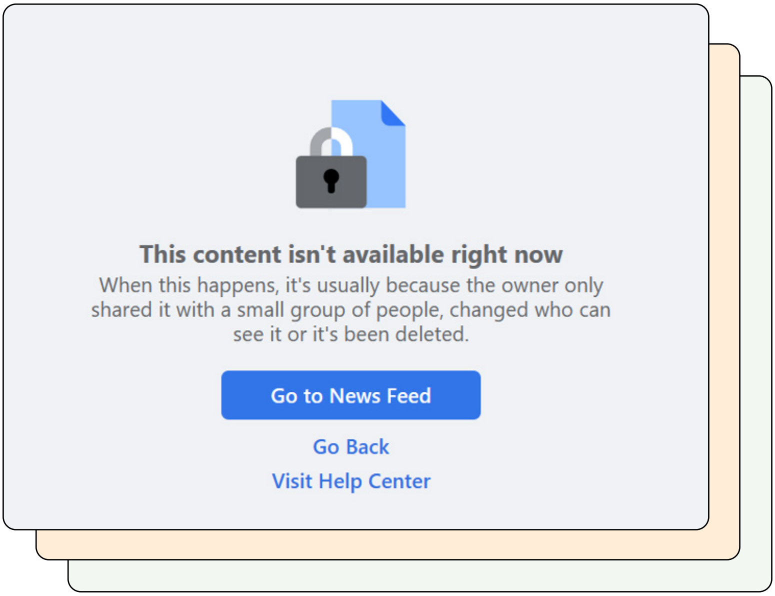
https://www.facebook.com/100084479785416/posts/142157408610251

https://www.facebook.com/100064560076979/videos/1362300001237651/

https://www.facebook.com/RadioIfm/videos/1516872728758062/

https://www.facebook.com/100066700361916/posts/503563795210295

https://www.facebook.com/100077772279822/posts/598772008736037



This content isn't available right now

When this happens, it's usually because the owner only shared it with a small group of people, changed who can see it or it's been deleted.

Screenshot 1: The Meta message indicating that content isn't available – Inaccessible links due to unavailability at the level of the content.

3. Another research limitation occurred when investigating Facebook pages account administrators locations. Due to a lack of information, the team cannot conclusively determine whether these pages were acquired, manipulated, or if their administrators are all located abroad. It is also plausible that VPNs or other methods were employed for specific reasons, further complicating the assessment of the pages' origins and administration.

General context of the data sample collected:

In the initial analysis, the data includes reactions to the new electoral law discussed in the context section. The collected content reveals a divided Tunisian social media landscape, with supporters and opponents expressing their perspectives on the law.

Supporters of the law argued that it is necessary to prevent individuals with criminal records from running for parliament. They view President Kais Saied's vision as revolutionary, and commend his efforts to combat corruption. On the other hand, opponents raise concerns that the law could be exploited to suppress critics of President Saied from participating in the elections. They also criticise the law for potentially weakening political parties and favoring wealthier candidates, by eliminating public campaign financing. Some opponents may associate President Saied with the term "Inqilab" (coup) in relation to the referendum and the new electoral law, indicating their refusal to support it.

Furthermore, the data analysis indicates that several major political parties announced they were boycotting the electoral process, deeming it fundamentally illegitimate.

Additional detailed analysis of the key findings is provided below.

Key Findings:

- **Key Finding 1: 1,330 posts out of 11,564 were annotated as “not verified” (11.5 per cent)**

  Out of the total 11,564 posts analysed, 1,330 (11.5 per cent) were classified as "not verified," indicating the presence of non-credible information, including mis/dis/malinformation or hate speech. Interestingly, these non-credible posts were shared by 58.4 per cent of the pages included in the data sample, suggesting that more than half of the monitored accounts (289 accounts) were involved in spreading disinformation during the elections.

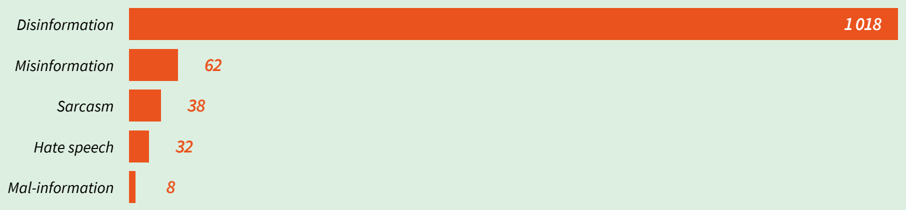


Figure 8: Distribution of the sub-classification of “not verified” content.

Further analysis revealed that the majority of the "not verified" content belonged to the “Photo” category, accounting for 54.81% of the content. Live videos accounted for 8.27 per cent of the non-credible content.

Link	288	21.65%
Live Video Complete	101	8.27%
Native Video	97	7.29%
Photo	729	54.81%
Status	94	7.07%
Video	3	0.23%
Youtube	9	0,86%
Total	**1 330**	**100.00%**

Figure 9: Type of “not verified” content by CrowdTangle

The analysis reveals that pages involved in sharing "not verified" content belong to various categories. The category "Politician" represented the highest percentage of pages sharing such content, accounting for 25.9 per cent of the total. Additionally, news sites contribute to 15.52 per cent of the "not verified" content, indicating that categories typically considered trustworthy sources are among the major contributors to the dissemination of unverified messages.

Politician	19.3%	25.90%
News_Site	14.0%	15.52%
Personal_Blog	11.0%	10.71%
Radio_Station	3.7%	7.42%
Activity_General	6.6%	5.51%
Media_News_Company	8.3%	4.84%
Community	5.1%	4.70%

Person	4.2%	4.53%
Blogger	4.7%	3.80%
Podcast	1.1%	2,31%
Broadcasting_Media_Production	1.0%	2.17%
Political_Party	2.4%	2.15%
Journalist	3.0%	1.73%
Political_Candidate	2.2%	1.44%
Media	0.8%	0.88%
Digital_Creator	1.0%	0.81%
Local	0.5%	0.78%
Artist	0.5%	0.70%
Tv_Channel	0.6%	0.55%
Community-Organization	2.0%	0.53%
Magazine	1.4%	0.51%
Topic_Food_Grocery	0.2%	0.45%
Advertising Marketing	0.1%	0.39%
NGO	1.3%	0.36%
Topic_Newspaper	0.8%	0.30%
City	0.1%	0.16%
Sports	0.1%	0.12%
Town_Hall	0.2%	0.11%
TV Show	1.4%	0.10%
Total	**100.0%**	**100.00%**

Figure 10: Categories of pages sharing "not verified" content, sorted by total interactions by CrowdTangle

- **Key Finding 2: 87.9 per cent of the non-verified content was classified as disinformation**

  The general context surrounding disinformation in the election period provided valuable insights into identifying these posts as falling under this heading. Through the analysis, we were able to identify the main techniques and narratives used in these campaigns during the elections; disinformation campaigns targeting political candidates used fake accounts, spreading conspiracy theories through re-direct links, and campaigns accusing candidates of "betrayal" -of the country- by receiving foreign funding.

- **Key Finding 3: 279 of the 477 pages shared non verified content – 58.4 per cent**

  Of the 477 pages examined, a total of 279 (58.4 per cent) shared non-verified content. This indicates a significant proportion of pages engaging in the dissemination of information that lacks proper verification. This finding raises concerns about the reliability and accuracy of the information being shared on these pages, and the potential impact on public perception and decision-making.

  The high percentage of pages sharing non-verified content suggests a widespread issue in terms of content validation and the dissemination of responsible information. It emphasises the need for improved fact-checking processes and critical evaluation of sources within the online ecosystem, particularly during sensitive periods, such as elections.

قهواجي البرلمان	46
Babnet	48
منية غزلاني خريف الصفحة الرسمية - Monia Ghozlani Khraief	41
TN اخبار	38
الراجل الي قهرهم	37
24 تونسنا	34
تحالف احرار	31
البلاغ	28
24 النهضة	26
المارد التونسي لتطهير الداخلية	25

Tuniscope	22
بوابة تونس - Tunigate	22
Kashf media - كشف ميديا	19
Diwan FM	17
المعز الحاج منصور	17
شبكة المدونين الاحرار	17
TAKRIZ	15
للحديث بقية	15
Tounesna FM	14
جريدة الحرية التونسية - AL Horria	14
الصفحة الرسمية لموقع النهار نويز	13
الوطنيون التوانسة - Les Nationalistes Tunisiens	13
بالسواك الحآر	13
Total	**1330**

Figure 11: Pages involved in disinformation campaigns (ranked by the most active pages)

Figure 11 presents the top pages involved in disinformation campaigns, ranked based on their level of activity, which is determined by the number of posts annotated as disinformation content. Among these pages are:

1. **"Kahweji Elbarlamen – قهواجي البرلمان"** – This is a public page, categorised as politician.
2. **"Babnet"** – This page belongs to a news media outlet.
1. **"Monia Ghozlani Khraief – منية غزلاني خريف الصفحة الرسمية"** – This is the official page of a candidate, and the page is categorised as a blogger.

Identifying those pages that shared non-verified content is a crucial step in understanding the scope and scale of the problem. By quantifying the number of pages involved, it becomes evident that a significant portion of the online landscape is susceptible to spreading potentially misleading or false information.

- **Key finding 4: 143 of 1,330 “not-verified” messages (10.7 per cent) directly targeted candidates**

An additional key finding reveals that 143 of the 1330 “not-verified” messages (10.7 per cent) were directly targeting candidates. These messages focused specifically on spreading potentially misleading or false information about candidates, representing a deliberate effort to influence public opinion and potentially harm the reputation or credibility of that candidate.

The remaining not-verified messages were found to contain disinformation about the election process in general, encompassing the main narratives mentioned earlier. This includes the dissemination of fake news, the weaponisation of mockery in mis-disinformation narratives, and the promotion of conspiracy theories. These disinformation campaigns aimed to create confusion, undermine trust in the electoral process, and manipulate public perceptions.

The large number of instances of the direct targeting of candidates and the broader dissemination of disinformation about the election process emphasises the need for increased scrutiny and critical evaluation of information during election periods. It highlights the potential impact that false or misleading narratives can have on shaping public opinion and influencing electoral outcomes.

فاطمة المسدي	**20**
محمد علي	**19**
رشدي الرويسي	**7**
هدى خليل	**7**
رمزي الشتوي	**4**
سجيعة الجلولي	**4**
فتحي رجب	**4**
محمد ماجدي	**4**
مهى عامر	**4**
نجلاء بنميلود	**4**
ياسين مامي	**4**
الطيب الطالبي	**3**

حمدي بن صالح	3
نورة الشبراك	3
الساسي علية	2
جلال خدمي	2
حاتم القلعي	2
سرحان الناصري	2
عماد أولاد جبريل	2
مكرم اللقام	2
نجلاء اللحياني	2
هالة جاب الله	2
وفاء بن سليمان	2
أحمد بنور	1
أحمد سعيداني	1
أسامه الولهازي	1
العربي قادري	1
أميرة شرف الدين	1
بسمة الهمامي	1
بوعلي رابح	1
بيرم أنيس القيلوزي	1
ثابت العابد	1
حسام بوقراص	1
حسن بوسامة	1
حنان بيبي	1

رضا دلاعي	1
رفيق بن دراه	1
رؤوف الفقيري	1
سرور محفوظي	1
سميره نصير	1
سيرين المرابط	1
شكري بن البحري	1
طارق الخوفي	1
طارق براهمي	1
عاطف بن حسين	1
عبدالحليم بوسمة	1
عبدالسلام الدحماني	1
عمار عيدودي	1
ماهر بوبكر الحضري	1
محمد أحمد	1
محمد العشي	1
محمد عمار	1
مريم الشريف	1
هالة الطرودي	1
هشام حسني	1
هناء الحداد	1
وليد بن جاوحد	1
يسرى الساسي	1
Total	**143**

Figure 12: Candidates targeted by disinformation campaigns (Sorted by the most-mentioned candidates)

– Key Finding 5: 45 per cent of the not-verified messages targeting candidates directly targeted women candidates

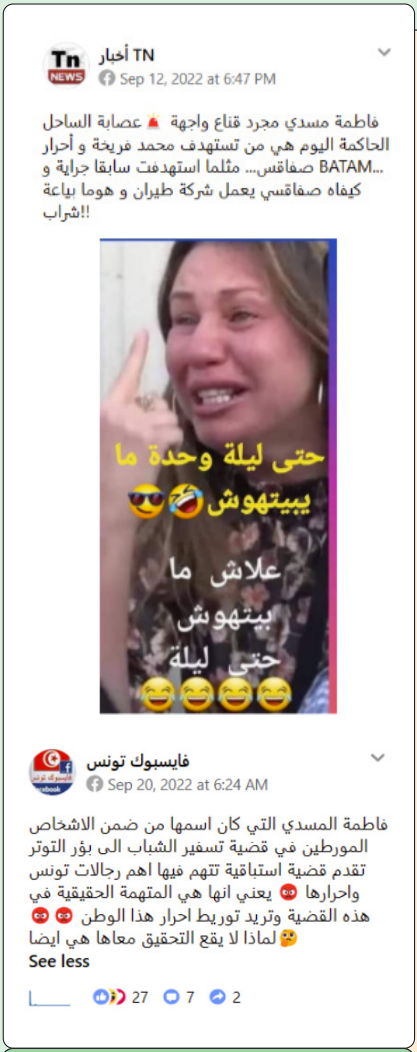
The data analysis reveals some concerning trends regarding the content targeting candidates in the context of an election. Out of the content directly aimed at candidates, which constitutes 10.7 per cent of the total content, a significant portion (45 per cent) was found to be spreading disinformation and false news about women candidates. This misinformation was aimed at undermining the credibility and trustworthiness of the individuals mentioned.

It is worth noting that women candidates made up only 11.5 per cent of the overall candidate pool. This low percentage already indicates limited participation of women in the electoral process. However, the fact that 45 per cent of the disinformation and hate speech messages are specifically targeting women candidates is a significant finding that warrants attention.

This key finding highlights not only the disproportionately high percentage of disinformation and hate speech aimed at women candidates, but also the potential challenges and obstacles they face in their electoral campaigns. The targeted dissemination of false news and negative messaging can have detrimental effects on the reputations and public perceptions of women candidates, further hindering their chances of success.

Examples:

1. Posts that allegedly target the candidate **Fatma Mseddi**, accusing her of involvement in the sending of young people to Syria, and then blaming other actors to avoid responsibility.
2. A post targeting candidate **Syrine Mrabet** shows that she was photographed with Kais Saied, Nabil Karoui, and Saief Eddine Makhlouf. The post stated "She lacks both "religion and dignity" and is excessively close to many influential individuals, which may lead to her automatic election."
3. Claims that a candidate named "Hanen Bibi" was involved in the theft of a coffee shop.
4. A post targeting Zakiya Kasraoui, a candidate who, despite having 5,000 followers on Instagram, was unable to advance to the second round. Furthermore, another physical insult was directed towards this candidate, calling her "Monkey" implying that she looks like one.
5. The post targeting candidate Mariem Laghmani. The post highlighted the left-affiliated candidate's habit of unserious behavior. Despite significant endorsements, she faced criticism for both her political alignment and negative interpersonal engagements. These actions likely influenced a negative perception that harmed her political reputation and re-election prospects.



أخبار TN
Sep 12, 2022 at 6:47 PM

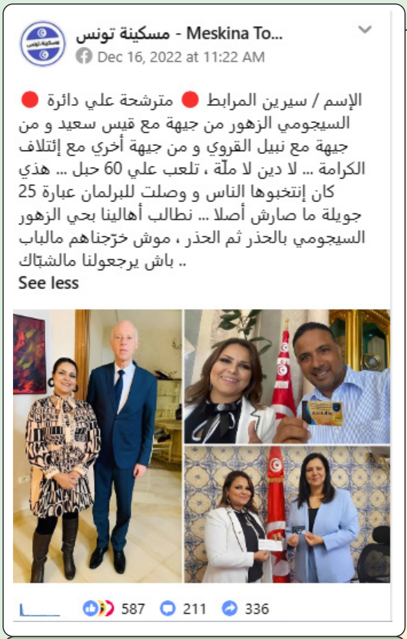
فاطمة مسدي مجرد قناع واجهة 🚨 عصابة الساحل الحاكمة اليوم هي من تستهدف محمد فريخة و أحرار صفاقس... مثلما استهدفت سابقا جراية و BATAM... كيفاه صفاقسي يعمل شركة طيران و هوما بياعة شراب!!

فايسبوك تونس
Sep 20, 2022 at 6:24 AM

فاطمة المسدي التي كان اسمها من ضمن الاشخاص المورطين في قضية تسفير الشباب الى بؤر التوتر تقدم قضية استباقية تتهم فيها اهم رجالات تونس واحرارها 😡 يعني انها هي المتهمة الحقيقية في هذه القضية وتريد توريط احرار هذا الوطن 😡😡 لماذا لا يقع التحقيق معاها هي ايضا 🤔
See less

27 7 2

1

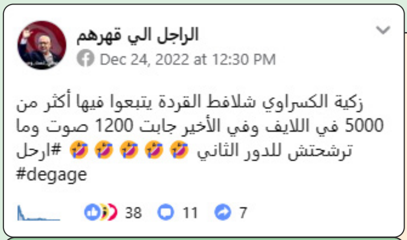


مسكينة تونس - Meskina To...
Dec 16, 2022 at 11:22 AM

الإسم / سيرين المرابط 🔴 مترشحة علي دائرة 🔴 السيجومي الزهور من جيهة مع قيس سعيد و من جيهة مع نبيل القروي و من جيهة أخري مع إئتلاف الكرامة ... لا دين لا ملّة ، تلعب علي 60 حبل ... هذي كان إنتخبوها الناس و وصلت للبرلمان عبارة 25 جويلة ما صارش أصلا ... نطالب أهالينا بحي الزهور السيجومي بالحذر ثم الحذر ، موش خرّجناهم مالباب باش يرجعولنا مالشبّاك ..
See less

587 211 336

2

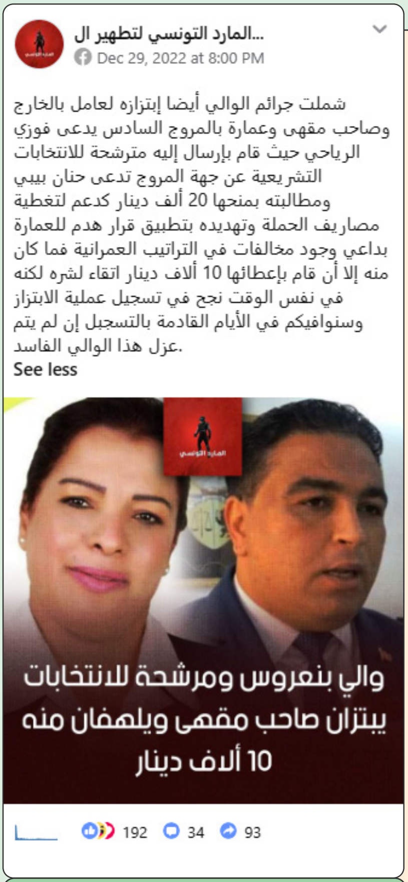


الراجل الي قهرهم
Dec 24, 2022 at 12:30 PM

زكية الكسراوي شلافط القردة يتبعوا فيها أكثر من 5000 في اللايف وفي الأخير جابت 1200 صوت وما ترشحتش للدور الثاني 🤣🤣🤣🤣🤣 #ارحل #degage

38 11 7

3



المارد التونسي لتطهير ال...

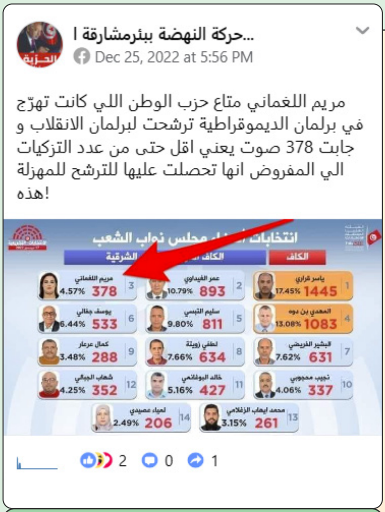
Dec 29, 2022 at 8:00 PM

شملت جرائم الوالي أيضا إبتزازه لعامل بالخارج وصاحب مقهى وعمارة بالمروج السادس يدعى فوزي الرياحي حيث قام بإرسال إليه مترشحة للانتخابات التشريعية عن جهة المروج تدعى حنان بيبي ومطالبته بمنحها 20 ألف دينار كدعم لتغطية مصاريف الحملة وتهديده بتطبيق قرار هدم للعمارة بداعي وجود مخالفات في التراتيب العمرانية فما كان منه إلا أن قام بإعطائها 10 ألاف دينار اتقاء لشره لكنه في نفس الوقت نجح في تسجيل عملية الابتزاز وسنوافيكم في الأيام القادمة بالتسجيل إن لم يتم عزل هذا الوالي الفاسد.

See less

192 34 93

4



حركة النهضة بيئرمشارقة ا...

Dec 25, 2022 at 5:56 PM

مريم اللغماني متاع حزب الوطن اللي كانت تهرّج في برلمان الديموقراطية ترشحت لبرلمان الانقلاب و جابت 378 صوت يعني اقل حتى من عدد التزكيات الي المفروض انها تحصلت عليها للترشح للمهزلة هذه!

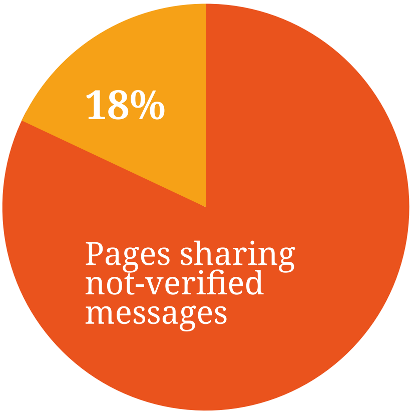
انتخابات أعضاء مجلس نواب الشعب
الكاف
الكاف الشرقية
1 ياسر قراري 17.45% 1445
2 عمر الفيداوي 10.79% 893
3 مريم اللغماني 4.57% 378
4 المهدي بن دوه 13.08% 1083
5 سليم التبسي 9.80% 811
6 يوسف جقالي 6.44% 533
7 البشير الفريضي 7.62% 631
8 لطفي زويتة 7.66% 634
9 كمال عرعار 3.48% 288
10 نجيب محجوبي 4.06% 337
11 خالد البوقلمي 5.16% 427
12 شهاب الجبالي 4.25% 352
13 محمد ايهاب الزغلامي 3.15% 261
14 لمياء عصيدي 2.49% 206

2 0 1

5

Screenshot 2: "Not verified" content (hate speech/ disinformation) about women candidates

– **Key finding 6:**

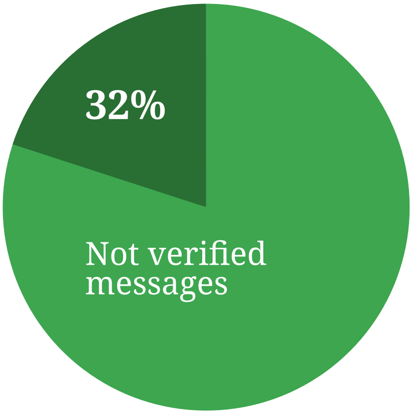


50 out of 279 pages (18%) were created within the election period – August 2022 to December 2022.

The finding highlights a concerning pattern during the election period, from August 2022 to December 2022. It indicates that a significant proportion of the analysed pages, 50 out of 279, or 18 per cent, were created within this timeframe. This observation raises suspicions that these pages may have been intentionally established to spread false news related to the elections, manipulate voter opinions, and create a sense of uncertainty among voters.

The deliberate creation of misleading pages to spread false information and manipulate public opinion is a significant problem that damages the credibility of elections. This can lead to various negative outcomes, such as influencing voters' behavior, distorting public discourse, and eroding trust in the democratic process. The existence of a significant number of such pages during election periods suggests targeted efforts to take advantage of the increased attention and susceptibility of voters at such times.

– Key Finding 7:



23 per cent of not-verified messages were published by pages not originating from Tunisia.

The given data suggests that out of all the content that had not been verified, 23 per cent was published by pages originating from countries other than Tunisia. These countries include Turkey (TR), France (FR), Indonesia (ID), Germany (DE.

This information implies that a significant portion of unverified messages comes from foreign sources, indicating the potential involvement of external actors or entities in disseminating such content. The higher percentage of not-verified messages originating from Tunisia itself could suggest a combination of both domestic misinformation and potentially genuine but unverified information circulating within the country. Further analysis and investigation would be needed to better understand the nature and impact of these messages.

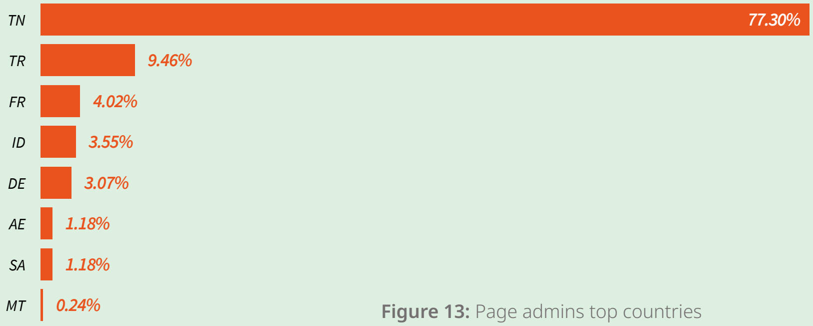
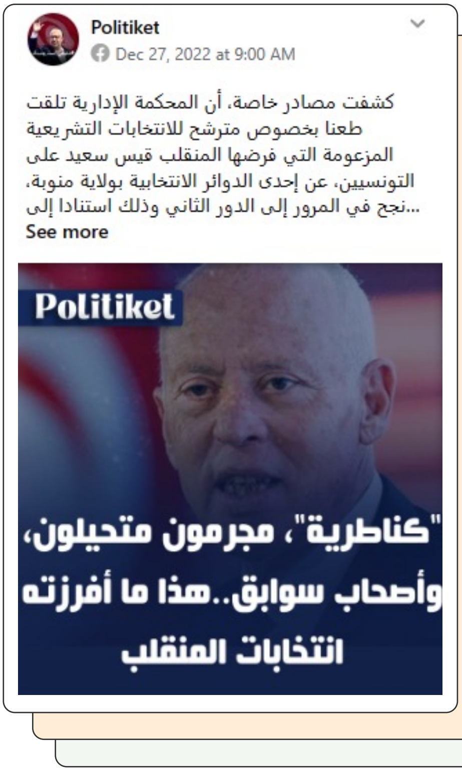


Figure 13: Page admins top countries

Narratives and trends of disinformation during the elections

Pages that engaged in disseminating false information around the elections focused their efforts on spreading false information about candidates. Disinformation campaigns often target political candidates by spreading false information about their personal lives, qualifications, or policy positions. These fabricated narratives aim to undermine the credibility and reputation of certain candidates, thereby influencing public perception and potentially swaying voter opinions.

Examples: Below are screenshots that display false news about candidates. One particular screenshot claims that a candidate named "Hanen Bibi" was involved in the theft of a coffee shop. Another screenshot describes a situation where, among the candidates, there are individuals with criminal records and histories of theft.

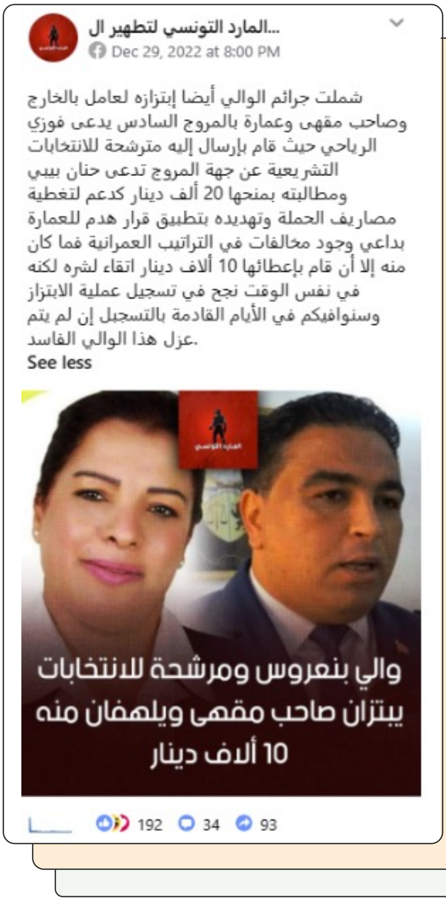


Politiket

Dec 27, 2022 at 9:00 AM

كشفت مصادر خاصة، أن المحكمة الإدارية تلقت
طعنا بخصوص مترشح للانتخابات التشريعية
المزعومة التي فرضها المنقلب قيس سعيد على
التونسيين، عن إحدى الدوائر الانتخابية بولاية منوبة،
نجح في المرور إلى الدور الثاني وذلك استنادا إلى...

See more



المارد التونسي لتطهير ال...

Dec 29, 2022 at 8:00 PM

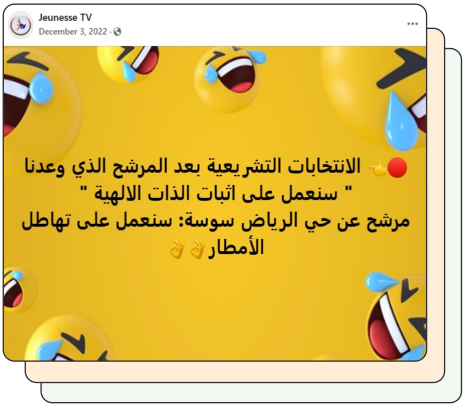
شملت جرائم الوالي أيضا إبتزازه لعامل بالخارج
وصاحب مقهى وعمارة بالمروج السادس يدعى فوزي
الرياحي حيث قام بإرسال إليه مترشحة للانتخابات
التشريعية عن جهة المروج تدعى حنان بيبي
ومطالبته بمنحها 20 ألف دينار كدعم لتغطية
مصاريف الحملة وتهديده بتطبيق قرار هدم للعمارة
بداعي وجود مخالفات في التراتيب العمرانية فما كان
منه إلا أن قام بإعطائها 10 ألاف دينار اتقاء لشره لكنه
في نفس الوقت نجح في تسجيل عملية الابتزاز
وسنوافيكم في الأيام القادمة بالتسجيل إن لم يتم
عزل هذا الوالي الفاسد.

See less

Screenshot 3: False information targeting candidates

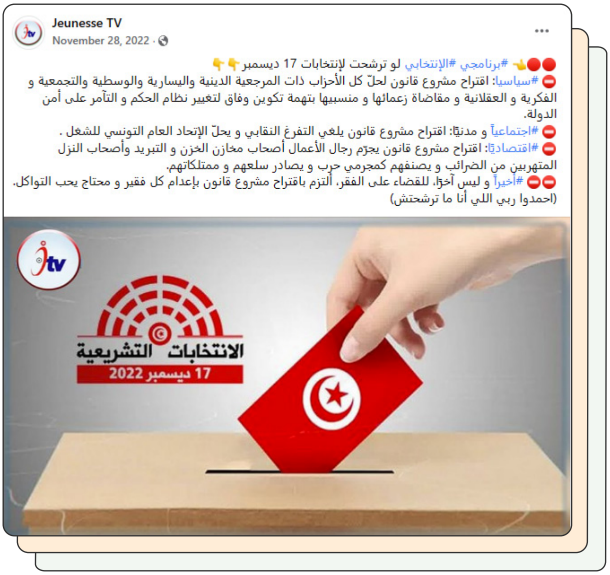
Another significant aspect was the weaponisation of mockery in mis-disinformation narratives. Disinformation campaigns frequently employ humor and satire as means to disseminate misleading information. By using mockery or satire, false narratives are disguised as jokes or memes, making them more likely to be shared and consumed uncritically. Such tactics can effectively blur the line between fact and fiction, making it challenging for individuals to distinguish between genuine news and disinformation.

Examples: In one screenshot, there is a campaign that mocks a candidate who claimed they would work on making it rain. The phrase was taken out of context, and manipulated in as sarcastic way. Another screenshot showcases sarcasm directed towards candidates in general, highlighting how they often make promises to achieve unrealistic and overly optimistic goals.Additionally, conspiracy theories were identified as key narratives within the disinformation content. These narratives aim to create confusion, sow distrust in established institutions, and manipulate public opinion to serve particular interests.



Jeunesse TV
December 3, 2022 ·

🔴👉 الانتخابات التشريعية بعد المرشح الذي وعدنا
" سنعمل على اثبات الذات الالهية "
مرشح عن حي الرياض سوسة: سنعمل على تهاطل
الأمطار👌👌

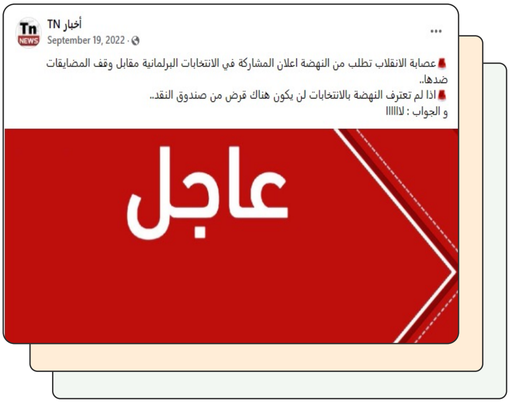


Jeunesse TV
November 28, 2022 ·

🔴🔴👉 #برنامجي #الإنتخابي لو ترشحت لإنتخابات 17 ديسمبر👇👇
⛔ #سياسيا: اقتراح مشروع قانون لحلّ كل الأحزاب ذات المرجعية الدينية واليسارية والوسطية والتجمعية و
الفكرية و العقلانية و مقاضاة زعمائها و منسبيها بتهمة تكوين وفاق لتغيير نظام الحكم و التآمر على أمن
الدولة.
⛔ #اجتماعياً و مدنيّا: اقتراح مشروع قانون يلغي التفرغ النقابي و يحلّ الإتحاد العام التونسي للشغل .
⛔ #اقتصاديّا: اقتراح مشروع قانون يجرّم رجال الأعمال أصحاب مخازن الخزن و التبريد وأصحاب النزل
المتهربين من الضرائب و يصنفهم كمجرمي حرب و يصادر سلعهم و ممتلكاتهم.
⛔⛔ #أخيراً و ليس آخرًا، للقضاء على الفقر، ألتزم باقتراح مشروع قانون بإعدام كل فقير و محتاج يحب التواكل.
(احمدوا ربي اللي أنا ما ترشحتش)

Screenshot 4: Examples of the weaponisation of mockery in mis/disinformation narratives

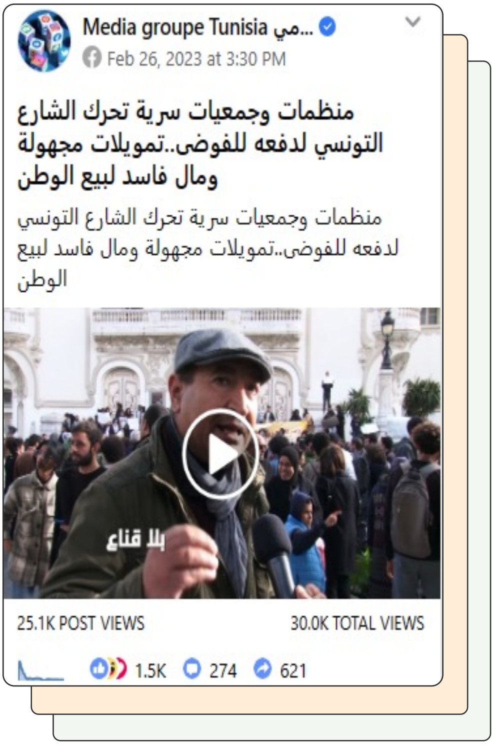
Examples: The screenshot below mentions demands for the participation of the Islamist conservative Ennahdha party in elections and to halt the arrests of its members, along with the International Monetary Fund's acceptance of a loan request for Tunisia.



أخبار TN
September 19, 2022 ·

عصابة الانقلاب تطلب من النهضة اعلان المشاركة في الانتخابات البرلمانية مقابل وقف المضايقات ضدها..
اذا لم تعترف النهضة بالانتخابات لن يكون هناك قرض من صندوق النقد..
و الجواب : لاااااا

Screenshot 5: Examples of conspiracy theories



Media groupe Tunisia مي...
Feb 26, 2023 at 3:30 PM

منظمات وجمعيات سرية تحرك الشارع التونسي لدفعه للفوضى..تمويلات مجهولة ومال فاسد لبيع الوطن

منظمات وجمعيات سرية تحرك الشارع التونسي لدفعه للفوضى..تمويلات مجهولة ومال فاسد لبيع الوطن

25.1K POST VIEWS 30.0K TOTAL VIEWS

1.5K 274 621

Screenshot 6: Examples of misleading content about foreign funds

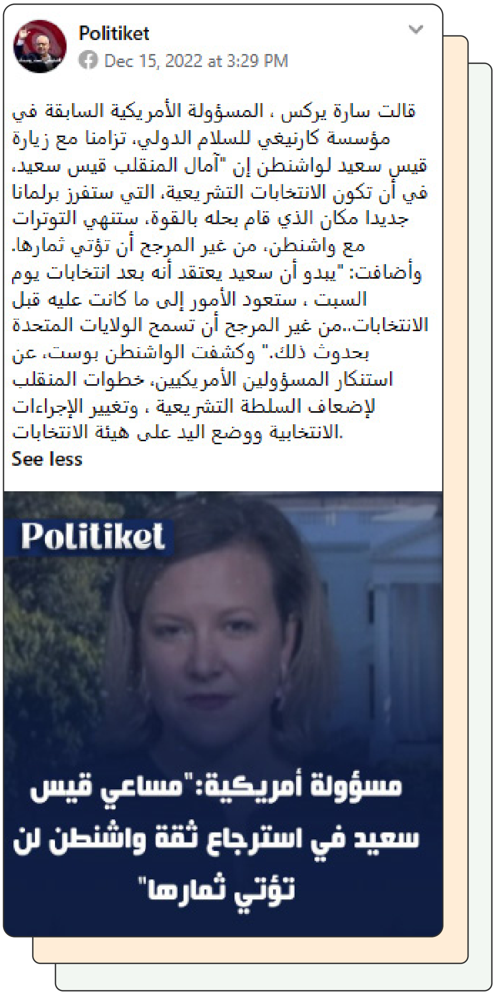
Another significant narrative observed in the disinformation posts during the election period was allegations of the involvement of foreign funds. This narrative aimed to create the perception that certain candidates were receiving financial support or interference from foreign entities, thereby casting doubt on their integrity and independence.

Examples: This post claims that "Organisations and secret funds" are deliberately inciting chaos in the Tunisian street. It further claims that these funds originate from foreign sources, remain anonymous, and involve corrupt money, with the intention of selling off the country.

The inclusion of this narrative within the disinformation content further underscored the deliberate nature of the misinformation campaigns. By alleging foreign involvement, the creators of the disinformation sought to exploit nationalist sentiments, fuel distrust in candidates or parties, and manipulate public opinion. This narrative played on existing fears and suspicions regarding external interference in domestic affairs, amplifying them for political gain.

At the same time, the suggestions that there was a lack of interest in the elections from other countries was also part of the disinformation content, also aimed at undermining the credibility and legitimacy of certain candidates, or the elections themselves. By implying that other nations did not consider the election worthy of support or attention, it sought to create a sense of doubt about the viability and legitimacy of the electoral process. This narrative aimed to erode trust in the election outcomes and raise questions about the fairness and integrity of the entire electoral system. By casting doubt on the interest or recognition of other countries, the disinformation campaign intended to weaken public confidence in the election process and, potentially, delegitimise the results.

Example: A statement from a former U.S. official named Sarah Yerkes asserts that Kais Saied's movement is insufficient to regain the confidence of Washington, which is partially false. The article, from the Washington Post, [5] does mention Yerkes, who studies Tunisia at the Carnegie Endowment for International Peace. She expresses skepticism as to whether the upcoming election and Saied's steps would resolve the strains between Tunisia and the United States. However, the article does not contain the phrase "Kais Saied's movement is insufficient to regain the confidence of Washington".



Politiket

Dec 15, 2022 at 3:29 PM

قالت سارة يركس ، المسؤولة الأمريكية السابقة في
مؤسسة كارنيغي للسلام الدولي، تزامنا مع زيارة
قيس سعيد لواشنطن إن "آمال المنقلب قيس سعيد،
في أن تكون الانتخابات التشريعية، التي ستفرز برلمانا
جديدا مكان الذي قام بحله بالقوة، ستنتهي التوترات
مع واشنطن، من غير المرجح أن تؤتي ثمارها.
وأضافت: "يبدو أن سعيد يعتقد أنه بعد انتخابات يوم
السبت ، ستعود الأمور إلى ما كانت عليه قبل
الانتخابات..من غير المرجح أن تسمح الولايات المتحدة
بحدوث ذلك." وكشفت الواشنطن بوست، عن
استنكار المسؤولين الأمريكيين، خطوات المنقلب
لإضعاف السلطة التشريعية ، وتغيير الإجراءات
الانتخابية ووضع اليد على هيئة الانتخابات.

See less

مسؤولة أمريكية:"مساعي قيس
سعيد في استرجاع ثقة واشنطن لن
تؤتي ثمارها"

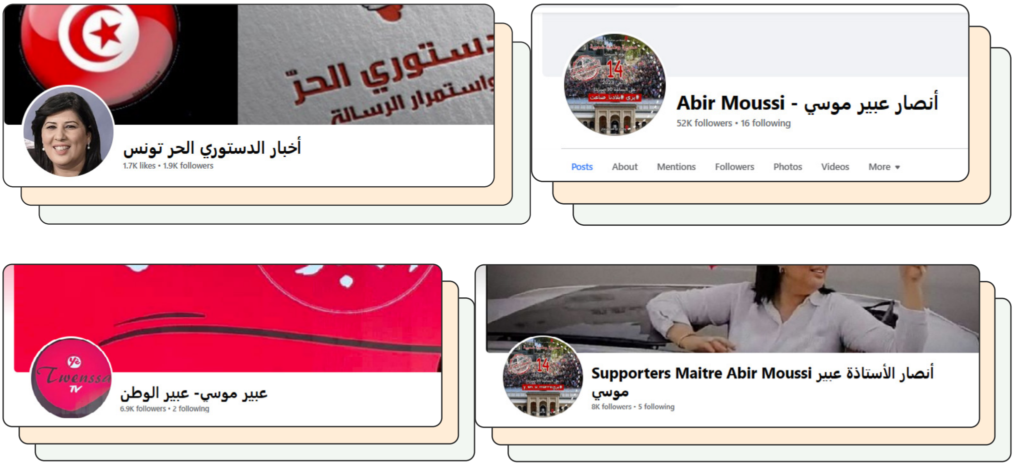
Screenshot 7: Examples of misleading content about a lack of interest in the elections from other countries.

Tactics and techniques used to publish disinformation during the elections

Fake pages and accounts of candidates:

During the parliamentary elections, fake accounts on Facebook were discovered by the Lab'Track team. These accounts claimed to be supporters of specific candidates by using the candidate's names in their page names. These were not, in reality, supporters of these candidates; their actual purpose was to spread disinformation about them.

[5] Missy Ryan, "Tunisia's Leader Defiantly Rejects U.S. Rebuke on Democratic Erosion", The Washington Post, 14 December 2011.



أخبار الدستوري الحر تونس

1.7K likes • 1.9K followers

Abir Moussi - أنصار عبير موسي

52K followers • 16 following

عبير موسي- عبير الوطن

6.9K followers • 2 following

Supporters Maitre Abir Moussi أنصار الأستاذة عبير موسي

8K followers • 5 following

Screenshot 8: Fake accounts supporting Abir Moussi involved in spreading “not verified” content

These fake accounts operate by disseminating false or misleading information about rival candidates. They may employ various tactics, including spreading rumors, creating fabricated stories, or distorting facts to manipulate public opinion. Their intention is to tarnish the reputation and credibility of opposing candidates, ultimately influencing voters' perceptions and decisions.

Deceptive Links:

Another concerning phenomenon identified by the data annotators was links that redirected users to web pages containing either disinformation or pages that are simply unavailable. These deceptive links played a manipulative role, by leading users to false or misleading content, or denying them access to legitimate information altogether.

These are often disguised as legitimate sources, such as news articles or reputable websites, enticing users to click on them. However, instead of providing accurate and reliable information, these links redirect users to web pages that propagate disinformation or lead to error pages, leaving users misinformed or frustrated.

Example link

Morphing pages:

Another concerning trend identified in the of Facebook pages was the phenomenon referred to as page morphing. This involved pages that underwent changes in their names, categories, or content types, particularly as the elections drew closer. Moreover, there were instances where these pages shared content that contradicted their designated categories or descriptions.

Page morphing, which was observed in the case of a significant number of monitored pages, carries significant risks, as it provides a platform for malicious individuals to spread disinformation and manipulate public opinion by altering the content and purpose of a page and, as a result, the original audience may be exposed to false or misleading information.

Morphing was used, in particular, as a tactic to target a wider audience, by converting popular pages in categories such as bloggers, artists, and digital creator magazines into propaganda platforms. This pattern has been consistently identified since 2019.

Through analysis, it was found that, out of 1,330 non-verified posts, 444 (33.4 per cent) originated from pages that are declared as belonging to categories not typically associated with political activity or displaying a nature of political involvement.

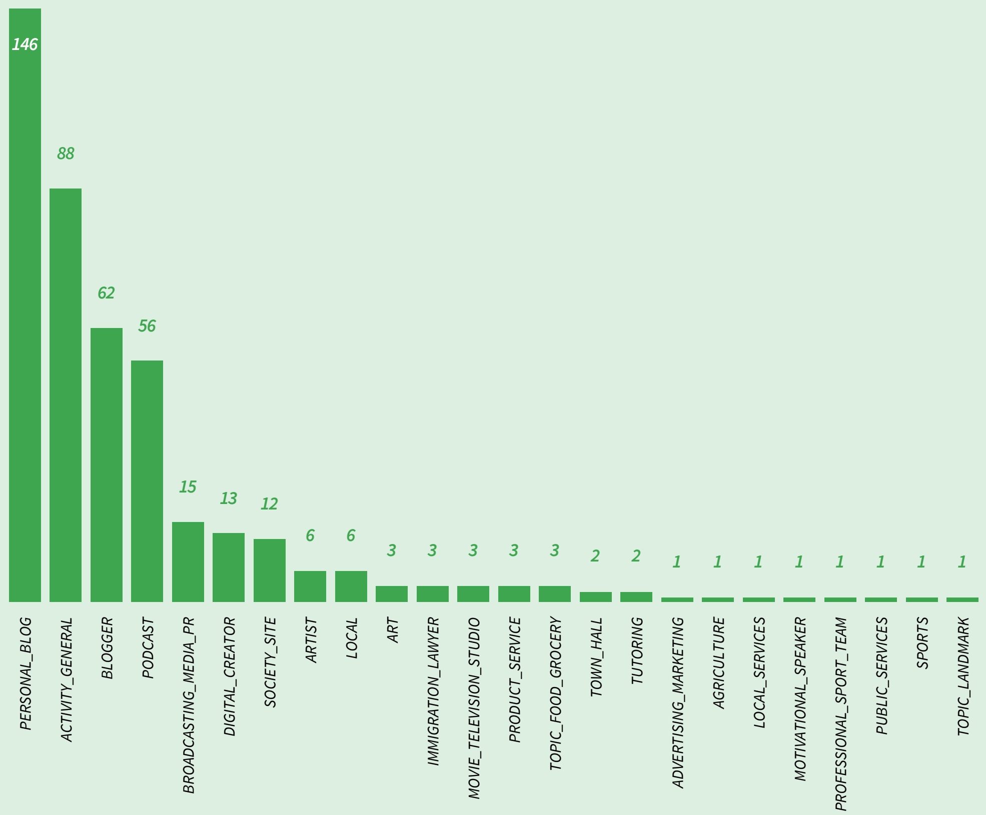


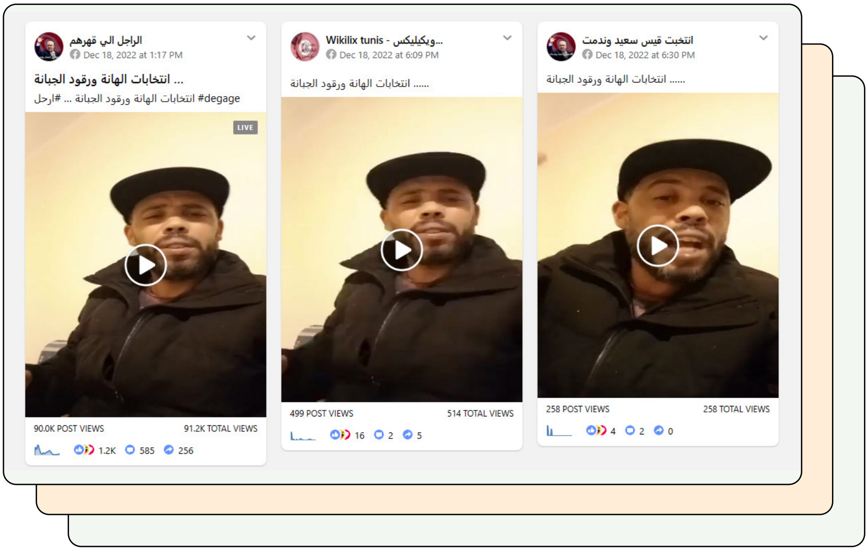
Figure 14: Categories of pages sharing not-verified content, created by CrowdTangle

Coordinated campaigns: Networks:

During the parliamentary elections, several Facebook page networks were identified as engaging in the strategic coordination of posting content and other social media activities. Their aim was to amplify specific political positions, The ultimate objective of these networks was to influence the political behavior of Tunisian citizens and shift the balance of the elections towards a desired outcome. The methodology section provides a detailed explanation of the steps and factors involved in identifying these networks.

Networks of pages were discovered sharing identical posts with the same format, writing style, and even video presentations. The coordinated nature of these pages extended to simultaneous live video broadcasts, indicating a deliberate effort to synchronise the dissemination of content. Notably, this behavior was observed previously in the second regional report published by DRI, related to online disinformation, hate speech, and the regional trends and local narratives in relation to the elections and suggesting a pattern of coordinated activity.

Examples of coordinated live videos:

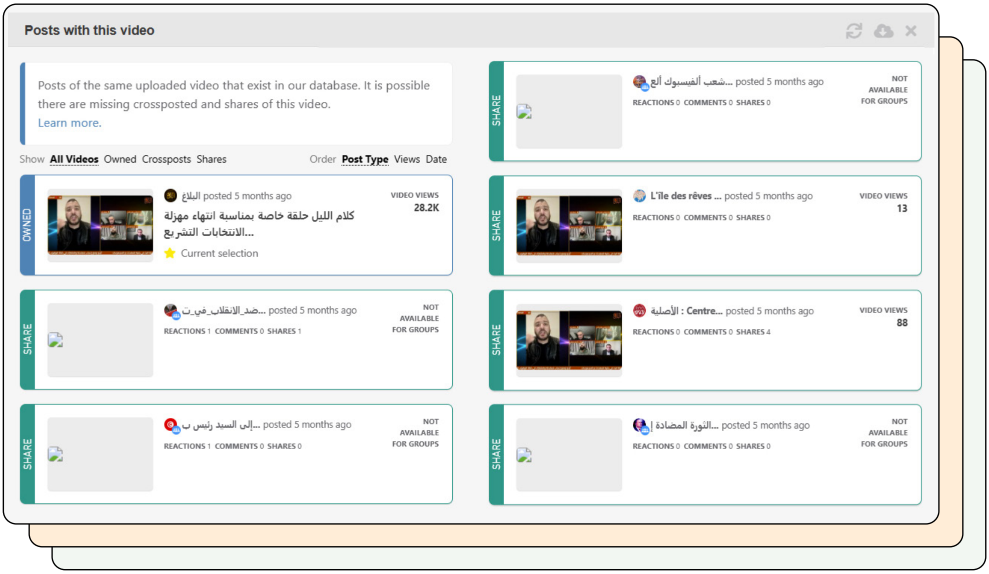


الراجل الي قهرهم
Dec 18, 2022 at 1:17 PM
انتخابات الهانة ورقود الجبانة ...
#degage انتخابات الهانة ورقود الجبانة ... #ارحل
LIVE
90.0K POST VIEWS 91.2K TOTAL VIEWS
1.2K 585 256

Wikilix tunis - ويكيليكس...
Dec 18, 2022 at 6:09 PM
انتخابات الهانة ورقود الجبانة ......
499 POST VIEWS 514 TOTAL VIEWS
16 2 5

انتخبت قيس سعيد وندمت
Dec 18, 2022 at 6:30 PM
انتخابات الهانة ورقود الجبانة ......
258 POST VIEWS 258 TOTAL VIEWS
4 2 0

Screenshot 9: Coordinated live video on Facebook pages



Posts with this video

Posts of the same uploaded video that exist in our database. It is possible there are missing crossposted and shares of this video. Learn more.

Show **All Videos** Owned Crossposts Shares Order **Post Type** Views Date

OWNED — البلاغ posted 5 months ago — كلام الليل حلقة خاصة بمناسبة انتهاء مهزلة الانتخابات التشريع... — ⭐ Current selection — VIDEO VIEWS 28.2K

SHARE — ضد_الانقلاب_في_ت... posted 5 months ago — REACTIONS 1 COMMENTS 0 SHARES 1 — NOT AVAILABLE FOR GROUPS

SHARE — إلى السيد رئيس ب... posted 5 months ago — REACTIONS 1 COMMENTS 0 SHARES 0 — NOT AVAILABLE FOR GROUPS

SHARE — شعب ألفيسبوك ألع... posted 5 months ago — REACTIONS 0 COMMENTS 0 SHARES 0 — NOT AVAILABLE FOR GROUPS

SHARE — L'île des rêves ... posted 5 months ago — REACTIONS 0 COMMENTS 0 SHARES 0 — VIDEO VIEWS 13

SHARE — الأصلية : Centre... posted 5 months ago — REACTIONS 0 COMMENTS 0 SHARES 4 — VIDEO VIEWS 88

SHARE — الثورة المضادة إ... posted 5 months ago — REACTIONS 0 COMMENTS 0 SHARES 0 — NOT AVAILABLE FOR GROUPS

Screenshot 10: Coordinated Live video identified by CrowdTangle*

- Based on the live video, CrowdTangle searches its database to find pages that have shared the same live video at the exact timestamp. Although many of these posts with live video have been deleted from the pages where they appeared, they still exist in the CrowdTangle database.

The content analysis revealed the presence of **three primary networks**, identified through the previously explained methodology:

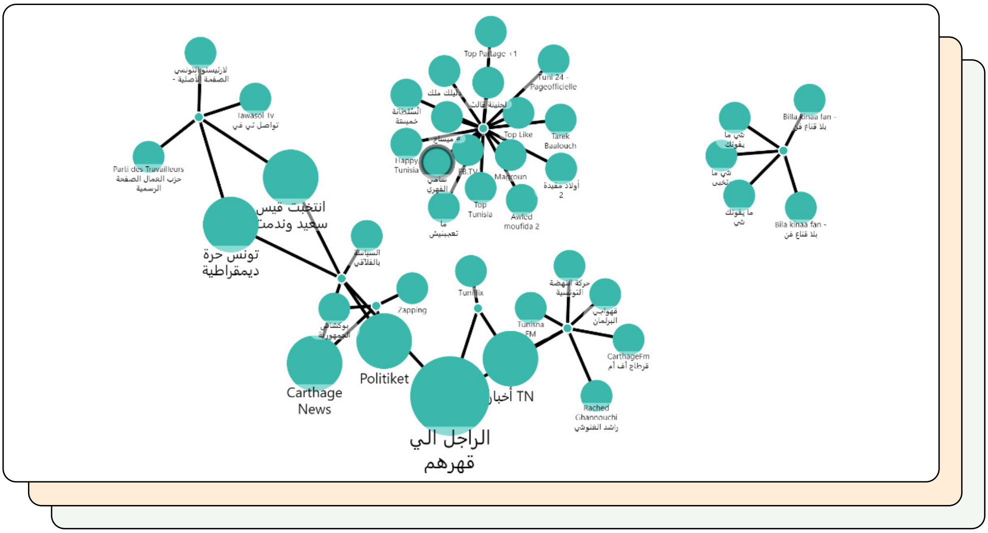
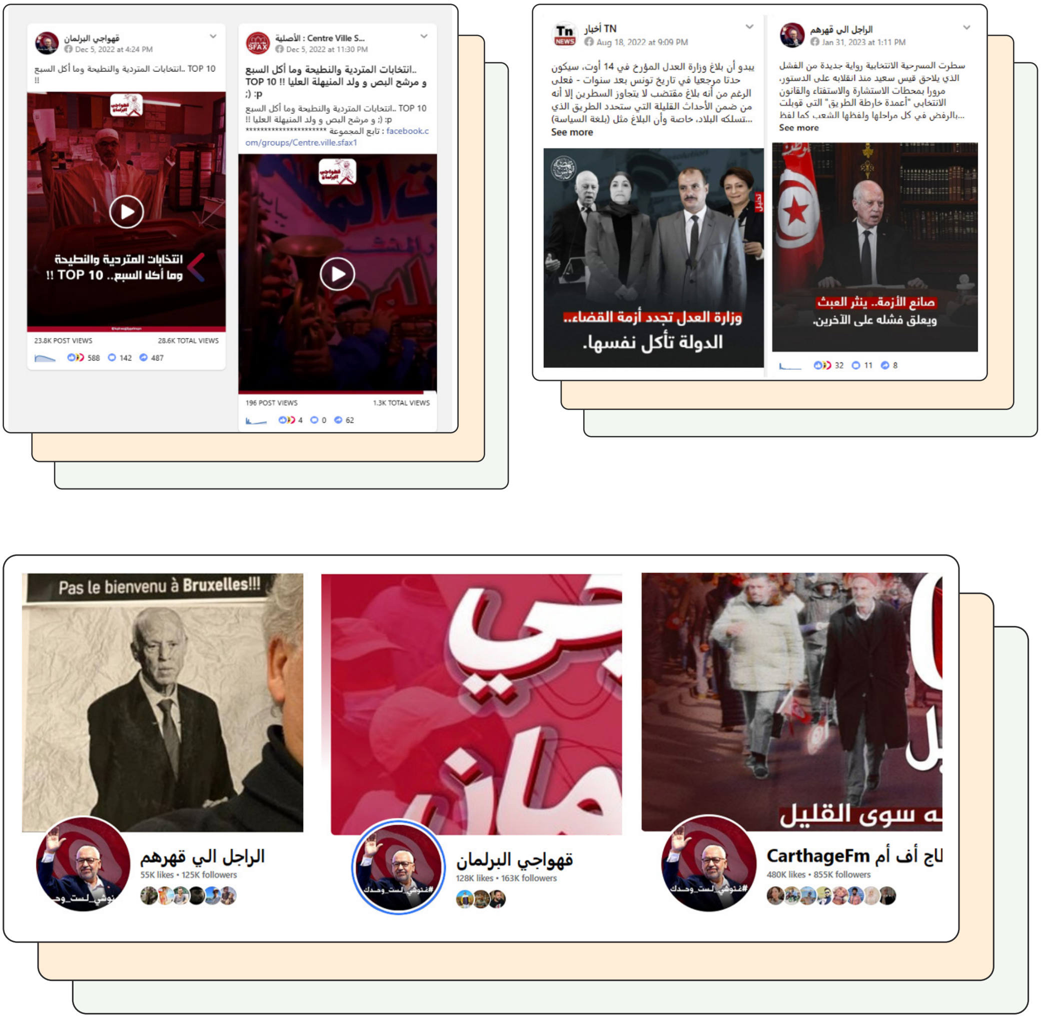


Figure 15: Networks detected by data analysis from the sample collected, created by Power BI

Network 1: The Anti-Kais Saied network:

In a deeper look at context shared by each network, the team found that the network containing the highest number of pages sharing not-verified content from mis/disinformation and hate speech was focused on attacks on President Saied.

This network displays a consistent pattern of sharing content that portrays Saied in a negative and hateful manner. The content describes him as corrupt, and refers to him as "INQILAB" (meaning “coup), while suggesting that he intends to destroy democracy in Tunisia.



قهواجي البرلمان
Dec 5, 2022 at 4:24 PM

TOP 10 ..انتخابات المتردية والنطيحة وما أكل السبع !!

23.8K POST VIEWS 28.6K TOTAL VIEWS
588 142 487

الأصلية : Centre Ville S...
Dec 5, 2022 at 11:30 PM

..انتخابات المتردية والنطيحة وما أكل السبع و مرشح البص و ولد المنيهلة العليا !! TOP 10
:p ;)

TOP 10 ..انتخابات المتردية والنطيحة وما أكل السبع و مرشح البص و ولد المنيهلة العليا !! :p ;)
********************* تابع المجموعة : facebook.com/groups/Centre.ville.sfax1

196 POST VIEWS 1.3K TOTAL VIEWS
4 0 62

أخبار TN
Aug 18, 2022 at 9:09 PM

يبدو أن بلاغ وزارة العدل المؤرخ في 14 أوت، سيكون حدثا مرجعيا في تاريخ تونس بعد سنوات - فعلى الرغم من أنه بلاغ مقتضب لا يتجاوز السطرين إلا أنه من ضمن الأحداث القليلة التي ستحدد الطريق الذي تسلكه البلاد، خاصة وأن البلاغ مثل (بلغة السياسة...
See more

الراجل الي قهرهم
Jan 31, 2023 at 1:11 PM

سطرت المسرحية الانتخابية رواية جديدة من الفشل الذي يلاحق قيس سعيد منذ انقلابه على الدستور، مرورا بمحطات الاستشارة والاستفتاء والقانون الانتخابى "أعمدة خارطة الطريق" التى قوبلت بالرفض في كل مراحلها ولفظها الشعب كما لفظ...
See more

32 11 8

الراجل الي قهرهم
55K likes • 125K followers

قهواجي البرلمان
128K likes • 163K followers

CarthageFm لاج أف أم
480K likes • 855K followers

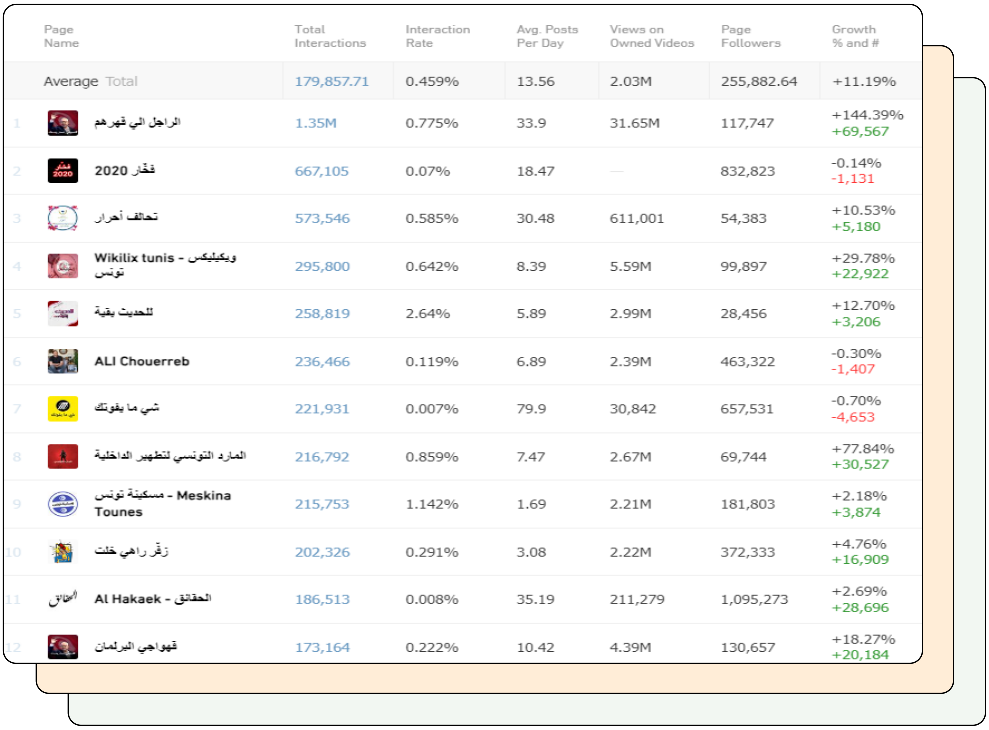
انتخابات المتردية والنطيحة
وما أكل السبع.. TOP 10 !!
وزارة العدل تجدد أزمة القضاء..
الدولة تأكل نفسها.
تحليل
صانع الأزمة.. ينثر العبث
ويعلق فشله على الآخرين.
Pas le bienvenu à Bruxelles!!!
ـه سوى القليل
#غنوشي_لست_وحدك

Screenshot 11: Pages involved in the network attacking President Saied

A significant finding by the team was that the pages within this network shared a common visual identity, indicating a coordinated effort to convey a cohesive message and establish a recognizable brand across various online platforms. Interestingly, the main narrative of the misleading information and anti-Saied hate speech content is closely tied to the Ennahdha party which had the parliamnetary majority before the 25 July coup and is considered the main opponent of President Saied. This correlation can be attributed to the fact that since, July 25th, with the series of arrests targeting Ennahdha members, the Ennahdha movement has emerged as one of the primary forces in opposition to Saied, alongside Abir Moussi, president of the Free Constitutional Party since 2016, and major opponent of Saied's.

Another noteworthy discovery was that many page administrators in this network have origins linked to Turkey. This raises the possibility that the attacks against the system and the changes initiated by the president might be organised or influenced by a foreign power, suggesting potential foreign intervention in Tunisian affairs.

Among the identified pages in this network, one specific page stands out as the most interacted-with during the election period. This page amassed a significant following, with over 125,284 people actively following it. Moreover, the page generated over 1.35 million interactions, reflecting the substantial engagement it garnered.



	Page Name	Total Interactions	Interaction Rate	Avg. Posts Per Day	Views on Owned Videos	Page Followers	Growth % and #
	Average Total	179,857.71	0.459%	13.56	2.03M	255,882.64	+11.19%
1	الراجل الي قهرهم	1.35M	0.775%	33.9	31.65M	117,747	+144.39% +69,567
2	2020 فخّار	667,105	0.07%	18.47	—	832,823	-0.14% -1,131
3	تحالف أحرار	573,546	0.585%	30.48	611,001	54,383	+10.53% +5,180
4	Wikilix tunis - ويكيليكس تونس	295,800	0.642%	8.39	5.59M	99,897	+29.78% +22,922
5	للحديث بقية	258,819	2.64%	5.89	2.99M	28,456	+12.70% +3,206
6	ALI Chouerreb	236,466	0.119%	6.89	2.39M	463,322	-0.30% -1,407
7	شي ما يفوتك	221,931	0.007%	79.9	30,842	657,531	-0.70% -4,653
8	المارد التونسي لتطهير الداخلية	216,792	0.859%	7.47	2.67M	69,744	+77.84% +30,527
9	مسكينة تونس - Meskina Tounes	215,753	1.142%	1.69	2.21M	181,803	+2.18% +3,874
10	زقّر راهي خلت	202,326	0.291%	3.08	2.22M	372,333	+4.76% +16,909
11	Al Hakaek - الحقائق	186,513	0.008%	35.19	211,279	1,095,273	+2.69% +28,696
12	قهواجي البرلمان	173,164	0.222%	10.42	4.39M	130,657	+18.27% +20,184

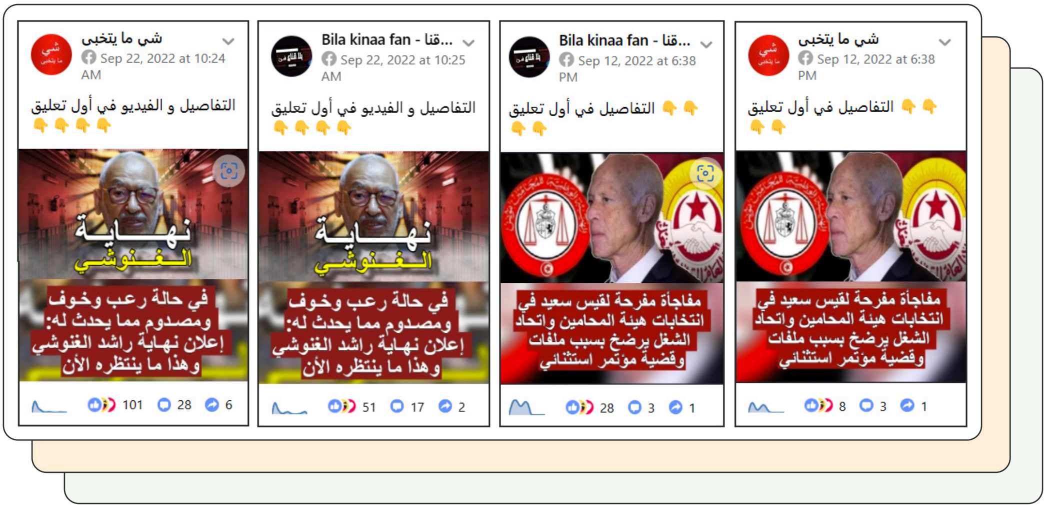
Figure 16: Leaderboard of the public pages with non-verified content.

Network 2: Supporters of President Kais Saied, and opposing the Ennahdha movement.

This network, consisting of five Facebook pages, specifically targeted the leaders of the Union Générale des Travailleurs Tunisiens - UGTT and Ennahda party, by spreading misleading and unfounded accusations. The posts shared on these pages were designed to create a sense of urgency and grab the attention of the audience by opening with the word "Urgent." This tactic was likely employed to generate immediate engagement and encourage users to read and share the content.

The analysis revealed that these pages were active between September 19 and October 5, during which they published an average of 3 posts per day.

Furthermore, the total interactions on these posts, including shares, likes, and comments, numbered 11,549. This indicates a relatively high level of engagement with the content.



شي ما يتخبى
Sep 22, 2022 at 10:24 AM
التفاصيل و الفيديو في أول تعليق 👇👇👇👇
نهاية الغنوشي
في حالة رعب وخوف ومصدوم مما يحدث له: إعلان نهاية راشد الغنوشي وهذا ما ينتظره الأن
101 28 6

Bila kinaa fan - قنا...
Sep 22, 2022 at 10:25 AM
التفاصيل و الفيديو في أول تعليق 👇👇👇👇
نهاية الغنوشي
في حالة رعب وخوف ومصدوم مما يحدث له: إعلان نهاية راشد الغنوشي وهذا ما ينتظره الأن
51 17 2

Bila kinaa fan - قنا...
Sep 12, 2022 at 6:38 PM
التفاصيل في أول تعليق 👇👇👇👇
مفاجأة مفرحة لقيس سعيد في انتخابات هيئة المحامين واتحاد الشغل يرضخ بسبب ملفات وقضية مؤتمر استثنائي
28 3 1

شي ما يتخبى
Sep 12, 2022 at 6:38 PM
التفاصيل في أول تعليق 👇👇👇👇
مفاجأة مفرحة لقيس سعيد في انتخابات هيئة المحامين واتحاد الشغل يرضخ بسبب ملفات وقضية مؤتمر استثنائي
8 3 1

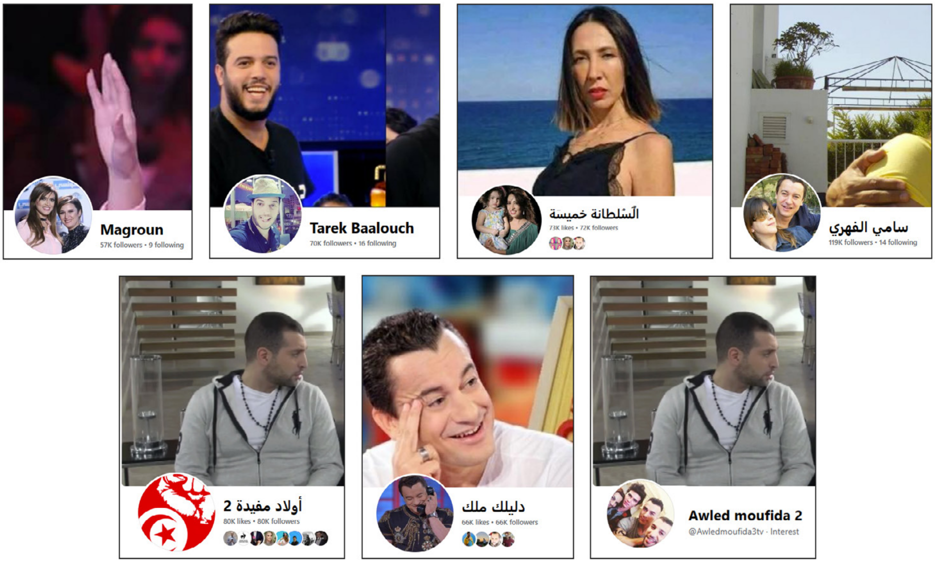
Screenshot 12: Pages involved in Network 2

Network 3: Mis/disinformation content and Sarcasm:

This network consists of 16 Facebook pages that are primarily focused on spreading unverified content about candidates and the elections in general. It is important to note that these pages do not take any particular position but, instead, engage in the dissemination of potentially misleading information, regardless of the target. The team observed that these pages are fake accounts impersonating famous Tunisian actors and animators, such as Sami Fehri, Samira Magroun, and Tarek Baalouch. These impersonations raise concern, as they can easily mislead their followers and manipulate opinions, due to the popularity and influence of these individuals.

Additionally, the team also noted the presence of fake pages related to popular Tunisian television series, such as "Awled Moufida."

These fake pages collectively amassed a total of 535,000 followers. The use of these pages, which portray themselves as belonging to popular figures and series, further contributes to the potential for misleading followers, manipulating opinions, and undermining of the democratic process as a whole.



Magroun
57K followers • 9 following

Tarek Baalouch
70K followers • 16 following

اَلسّلطانة خميسة
73K likes • 72K followers

سامي الفهري
119K followers • 14 following

أولاد مفيدة 2
80K likes • 80K followers

دليلك ملك
66K likes • 66K followers

Awled moufida 2
@Awledmoufida3tv · Interest

Screenshot 13: Pages involved in Network 3

Conclusion

The report highlights the impact of disinformation and misinformation on public opinion and electoral outcomes during Tunisia's parliamentary elections. It points out that the amendments made to the electoral law through presidential decree 55 were criticised for neglecting political parties, removing quotas for women and youth, and lacking equal geographic representation. As a result, boycott campaigns emerged, questioning the legitimacy of the first-ever individual-based candidate election.

The analysis focused on 477 monitored Facebook pages during the pre-election, election, and post-election periods. It revealed the utilisation of various tactics, of varying levels of intensity, to spread disinformation and hate speech employed by different factions with political interests. The report also explored how the public reacted to these tactics while the competition for power in the legislative branch was ongoing. It observed trends such as morphing pages, coordinated networks, decentralised management, and sarcastic campaigning

Comparisons were drawn with the 2019 presidential elections and the constitutional referendum, indicating similar patterns of coordinated live videos, morphing and networked pages spreading misleading content and hate speech. This suggests an ongoing pattern.

One significant observation was that of targeted attacks against women candidates, with individuals being targeted based on their appearance, gender, and political affiliation. These findings emphasise the challenges faced by candidates, and particularly women, during the electoral process.

The report underscores the influence of disinformation and hate speech on Tunisia's parliamentary elections, highlighting specific issues related to the electoral law amendments, patterns of deceptive tactics, and targeted attacks against candidates.

Recommendations

The study's findings on morphing pages, memes, and multi-country managed pages used for manipulative purposes highlight the need for increased regulation and oversight. One possible approach would be for social media platforms to invest greater resources in the moderation of content in Arabic, and in greater technical capabilities to monitor harmful content in non-English contexts. Platforms should regularly audit and monitor their pages, assess their platforms' vulnerability to online manipulation prior to election cycles (e.g., publishing country factsheets), conduct transparent investigations into any reports of manipulative behavior, and share their findings with trusted regional stakeholders. Another recommendation is that CSOs become more active in raising public awareness and providing education on the risks and potential harms of disinformation spread on social media.

Targeted awareness campaigns and public service announcements should be developed to highlight the importance of critical thinking, fact-checking, and media literacy skills.

Social media companies should also play a role in this, by investing in pre-bunking campaigns and providing users with more information on where they can find verified information about the election cycle, and how to identify and avoid fake or misleading content.

Annex 1: Keyword list

Arabic	English
تزوير النتائج	Forgery of results
طعون	Appeals
النتائج	Results
النتائج النهائية للاستفتاء	Final referendum results
الفصل 139	Article 139
دستور جديد	New constitution
قانون انتخابات سعيد	Saied's election law
قانون الانتخابات	Election law
تنقيح قانون الانتخابات	Amendment of the election law
اجراءات الترشح	Candidacy procedures
قواعد الترشح	Candidacy rules
إعلان الهيئة العليا المستقلة للانتخابات	Announcement of the Independent High Election Commission
فاروق بوعسكر	Farouk Bouasker
ترشحات الانتخابات التشريعية	Legislative election candidacies
تقديم الترشحات	Candidacy submissions

مرشح	Candidates
تزكيات	Endorsements
عدد الترشحات الجملي	Total number of candidacies
تمديد فترة قبول الترشحات	Extension of the candidacy acceptance period
قائمات الناخبين النهائية	Final electoral lists
إعلام المشاركين	Participant notification
انقضاء الطعون	Expiration of appeals
الحملة الانتخابية	Election campaign
المترشحين	Candidates
حملة انتخابات ديسمبر	December election campaign
الدوائر الانتخابية	Electoral districts
حملة مضللة	Misleading campaign
انتهاء الحملة في الخارج	End of the campaign abroad
خرق الحملة الانتخابية	Violation of the election campaign
ضعف الحملة الانتخابية	Weak election campaign
انتخابات ديسمبر ٢٠٢٢	December 2022 elections
الصمت الانتخابي	Election silence
خروقات	Violations
جرائم انتخابية	Election crimes
عزوف	Voter abstention
انتخاب	Voting
دوائر بالخارج	Overseas constituencies
اقبال ضعيف	Low voter turnout

نسب الاقبال	Voter turnout percentages
مترشح	Candidate
الهيئة العليا المستقلة للانتخابات	Independent High Election Commission
انتخاب	Election
خرق الصمت الانتخابي	Violation of campaign silence
نتائج الانتخابات	Election results
انتخابات تونس ٢٠٢٢	Tunisia 2022 elections
نتائج ضعيفة	Weak results
حملة انتخابات الدور الثاني	Second round election campaign
حصص التعبير المباشر	Direct expression sessions
انتهاء الحملة	End of the campaign
قائمات الناخبين النهائية	Final voter lists

2. Jordan Open Source Association (JOSA)

Nuha | An AI model to Detect Online Gender Based Violence against Women in Jordan

About the Jordan Open Source Association

The Jordan Open Source Association (JOSA) is a non-profit organisation based in Amman, Jordan, committed to promoting the principles of openness and accessibility in technology. JOSA's mission is centered on the belief that non-personal information should be freely accessible to all, in the form of open-source software. JOSA is also a strong advocate for the protection of personal information and the establishment of legal and technological frameworks that safeguard users' digital rights.

Through its initiatives, JOSA strives to promote a safe and inclusive digital environment for women, as gender-based cyberattacks are a major issue in the country.

Introduction

JOSA is developing an artificial intelligence (AI) model called “Nuha” (from the Arabic word for “mind”, or "brain"), which aims to detect hate speech against women in Jordanian digital spaces, such as social media platforms. In Jordan, the phenomenon of online gender-based violence (OGBV) is still under-researched, according to Siren Associates' 2021 Annual Report. [6] To ensure the active involvement of women in Jordanian digital public spaces, it is essential to analyse and study the discourses spread on these platforms, as well as the content directed at women through these platforms in a systematic and scientific manner. The processes of monitoring and studying these discourses are of a high degree of complexity, however, due to the enormous amount of harmful content against women on social media platforms, according to the Siren Associates report.

To contribute to this process, under the "Words Matter" project, the scope of Nuha focuses **specifically on detecting OGBV against women in Jordan's digital public spheres.**

JOSA's work focuses on the detection and classification of gender-based hate speech against women in Jordan. This encompasses a wide range of behaviors, from misogynistic insults to threats of violence, and can have serious consequences for the safety and well-being of women who are targeted by these forms of speech.

As Nuha is still under development, in this report, JOSA will provide insights into the data used to train the AI model and the methodology followed during both the research phase and the development of the Nuha model. This includes the selection process for the datasets sample, the use of machine-learning algorithms to develop the model, and the implementation of testing and validation procedures. In addition, the report will provide an oversight of the initial analysis of hate speech against women in Jordan, based on the dataset sample, including numbers and statistics that offer a clearer picture of data used to train Nuha and a glimpse into the situation regarding misogynist hate speech in the country. It also reflects the limitations and difficulties faced throughout the data collection and social media monitoring, the prevalence of different types of hate speech in various digital spaces, the factors that contribute to the spread of hate speech online, and the development of the tool . The report concludes by offering recommendations for social media platforms and relevant Jordanian authorities and non-profit organisations to combat cyber gender-based violence and sexist hate speech in digital spaces.

JOSA believes that the research and the development of Nuha offer an important contribution to ongoing efforts to promote gender equality and combat hate speech against women in Jordan, other Arabic-speaking countries, and beyond. This work is intended to inspire future research to produce gender-inclusive projects and tools.

Context

With the significant spread of social media platforms in recent years, interest in women's issues in digital spheres has begun to increase in Jordan.

[6] Siren Associates, "Annual Report", May 2021.

Notably, several organisations have emerged as key contributors to the study of this phenomenon at a national level, including, but not limited to the Sisterhood is Global Institute (SIGI - Tadamun), the SecDev Foundation, and the Jordanian National Commission for Women (JNCW).

However, despite the ongoing discussions and debates at the national level about integrating and empowering Jordanian women in the public and political spheres and making digital spaces in Jordan safer for women, an integrated system to protect women from OGBV is yet to be developed, as the current relevant legislation still falls short in protecting women in Jordan. For instance, according to the JNCW, the vast majority of women who face OGBV do not report it to the competent authorities (i.e., police and the Anti-Cyber Crimes Unit) due to a number reasons, such as social stigmatisation, bureaucratic procedures, overall lack of trust in public institutions in handling sensitive cases such as OGBV, and an unsecure process resulting from litigation procedures that require a personal presence for reporting. Therefore, only 28.7 per cent of the cases reported by women were solved.[7].

The existing legal and institutional systems and frameworks also fail to consider the social and cultural context specific to Jordanian women. This lack of consideration for women's reality in Jordanian society has severe consequences. For example, some forms of OGBV may lead to honor killings, where women are killed by family members. This grim reality makes women hesitant to officially report the digital threats they face, as doing so may result in severe social consequences, or threats to their health and lives. The fear of these consequences further discourages women from seeking help from official agencies to address cases of OGBV.[8]

In addition to the legal shortcomings, a large percentage of employees working in the relevant official authorities do not have sufficient knowledge to deal with such cases, due to a lack of education and training on gender concepts and related issues. According to a study by the United Nations High Commissioner for Refugees (UNHCR) [9], which took East Amman as its study sample, many service providers for women stated that they often did not know how to deal with cases of gender-based violence.

All of the above makes social media platforms in Jordan an unsafe space for women, given the amount of harassment they may face on these platforms, especially if they are involved as political or human rights activists.

Data and Methodology

JOSA adopted a machine learning lifecycle that consists of four phases – discovery, training, deployment, and monitoring.

In the discovery phase, the problem is defined, data is gathered, and goals are set. The training phase involves building and training the model, using the collected and classified data. Deployment integrates the model into the production

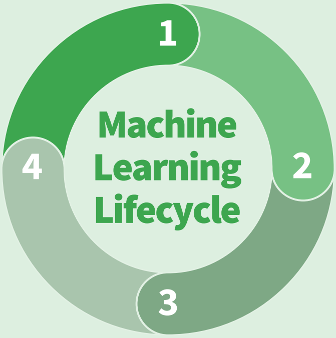
[7] https://women.jo/~women/sites/default/files/06-2022/%D%8AF%D%8B%1D%8A%7D%8B%3D%8A[...]%A%7D%20%85%9D%88%9D%8A%7D%84%9D%8B%3D8%9A%D%8A%7D%8B%3D8%9A.pdf

[8] Jordan Open Source Association (2022), Gender-Based Violence Against Women in Jordan: A policy Paper.

[9] GBV Sub Working Group, Jordan "Gender-Based Violence Risk Assessment for East Amman, October 2021", 1 November 2021.

environment. Monitoring continuously assesses the model's performance and makes necessary adjustments. Up until the reporting period, and by following this lifecycle, JOSA developed a Nuha model with a 72 per cent F1 score. As Nuha is still in the developing phase, JOSA is working on enhancing the F1 score, aiming for an F1 of 90-95 percent accuracy.

The below chart 01 summarises the machine learning lifecycle.



1. Discovery
Gather | Explore | Understand

2. Training
Experiment | Train | Validate

3. Deployment
Serve model | Score batches | Integrate

4. Monitoring
Log | Visualize | Alert

Chart 1: The Machine Learning Lifecycle

In order to start the machine learning lifecycle, JOSA established a methodology that explains the process to be adopted in building the AI model of Nuha. The methodology discussed two main aspects; the research part, which plays a key role in developing Nuha, and the technical part, that addresses definitions for non-technical readers.

Discovery Phase

Throughout the first phase (discovery), JOSA conducted multiple roundtables with the Women Digital Safety Alliance [10] on the machine learning lifecycle. Specifically, this phase included consultations with our TAMAM partners where we asked them to pinpoint

[10] The TAMAM Coalition is a women's alliance for digital security and a local Jordanian coalition that JOSA established as a partnership of local women rights organisations. The main goal of the coalition is to raise awareness of digital security and safety amongst women in Jordan.

women activists' accounts in Jordan, as well as to help us study the sociopolitical context of the status of online violence in the country. to help us study the sociopolitical context of the status of online violence in the country.

The research aspect focuses on the identification of accounts, monitoring, data collection, labeling, documentation, tools, and research in collaboration with TAMAM coalition. JOSA completed in-house desk research in order to identify a sample of social media accounts for women in Jordan that are subjected to online gender-based hate speech. We selected 83 Women in the Jordanian public spaces from different backgrounds and personas (e.g., influencers, journalists, human rights defenders, and politicians) presented in the chart 03. Approximately 50 per cent of the selected women's accounts belong to politicians or activists. The selected women's accounts were chosen based on their activity on social media, targets of online gender-based hate speech, or potential targets. In addition, 20 online Women Rights movements campaigns and hashtags that advocate for women's rights and civil and political rights for Jordanian women were selected to be part of the discovery phase.

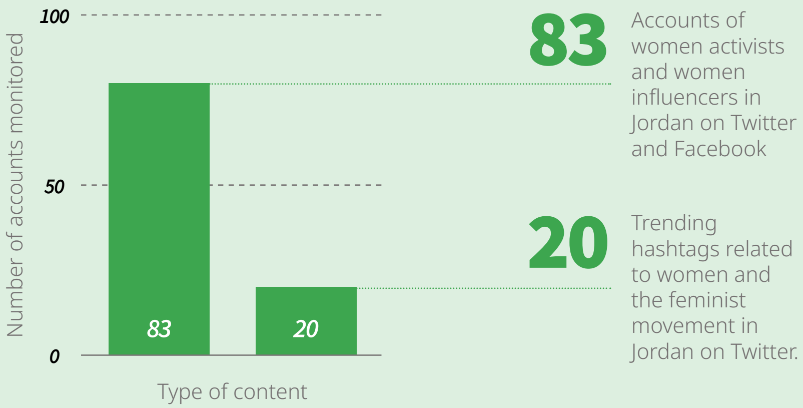


Chart 2: Sample breakdown

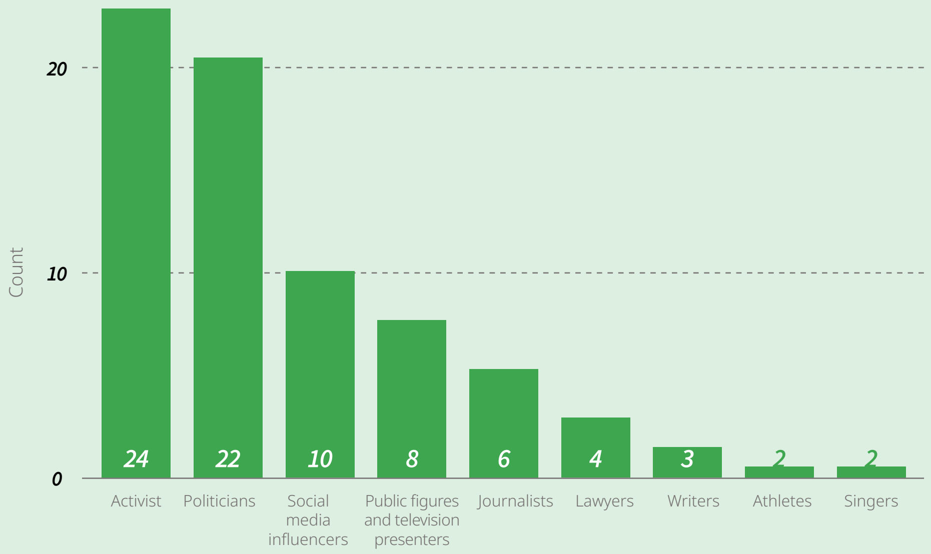


Chart 3: Sample Represented by Social Media Activity

In parallel, the TAMAM coalition played a key role in identifying the classification levels of data that should be used during the monitoring phase and data annotation. As a result, a mind-map of classification "annotation" with clear definitions was created. Three layers of classification were identified – 1) gender-based hate speech or not hate speech, to be used as an input for Nuha's development; 2) gender-based hate speech classification, to provide support for research and in identifying patterns; and 3) the basis of gender-based hate speech, to provide a better understanding of the context and to dive deep into what the hate speech is based on for example: hate speech based on socio-economic class, or race, or gender. Figure 01 reflects the three levels and the sub-classifications.

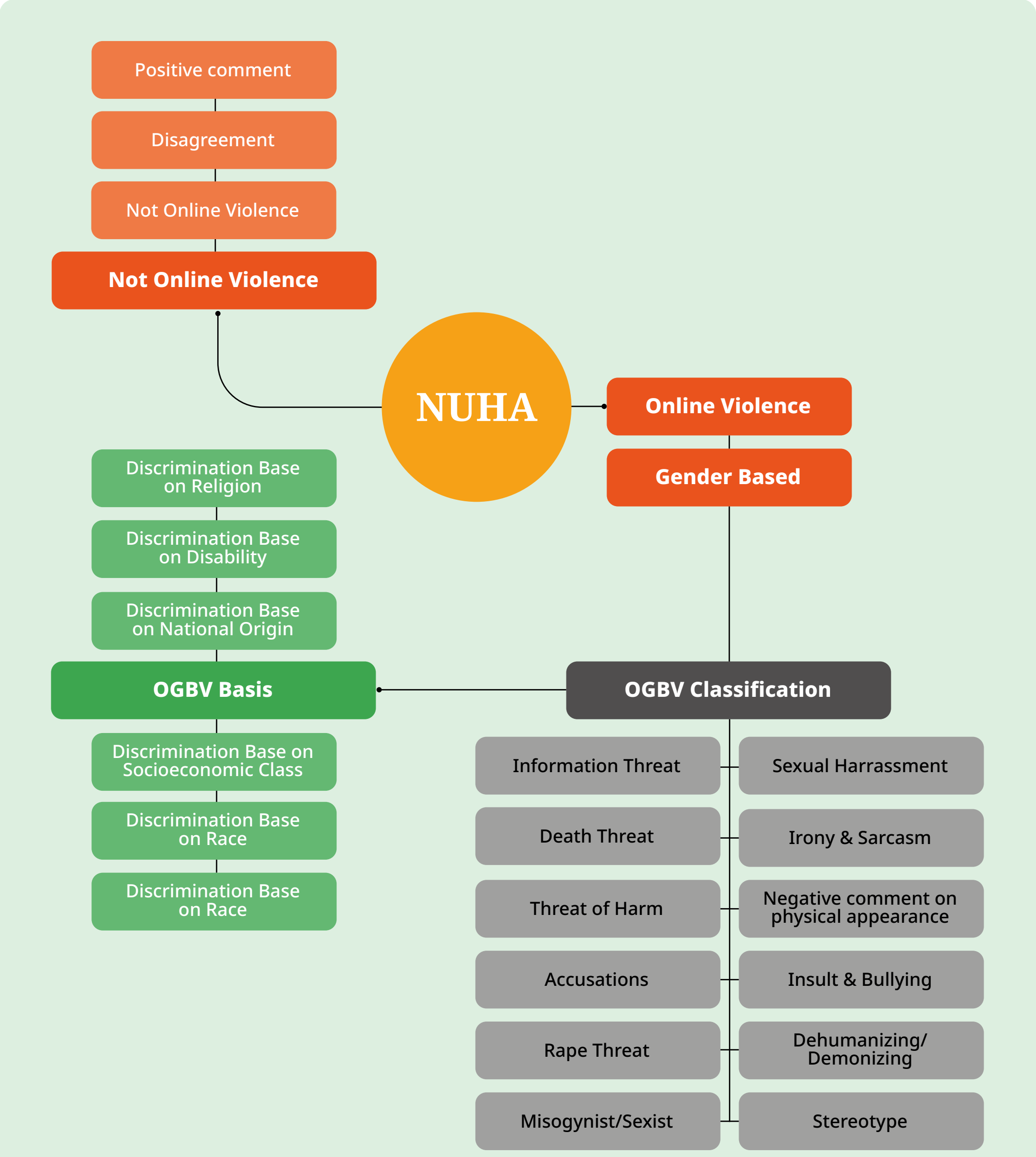


Figure 01: The Nuha Mind-map

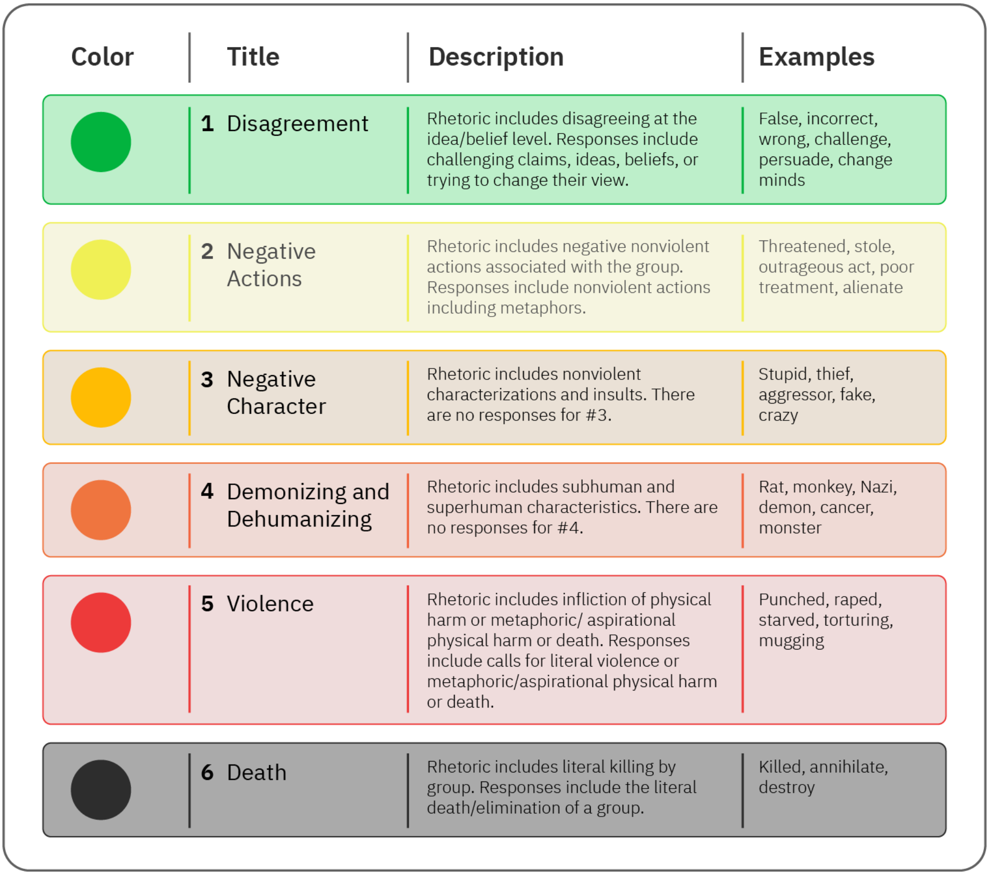
As reflected in Figure 01, the first layer labelling/classification of the mind-map consists of online violence and different categories on non-online violence . The reason for the different classification for the non-online violence is directly linked to the data required in developing the tool, to enable it to differentiate between hateful content, disagreement, and not hateful content. In simpler words, we cannot learn the color blue if we were not taught the other colors as well.

In addition, JOSA and the TAMAM coalition added clear definitions for some of the concepts that might cause confusion and difficulty during the annotation process, based on the context and situation in Jordan. Worth highlighting, the term “hate speech” is used in some laws in Jordan and has a particular definition (sometimes used to silence political activists), so our partner organisations in the country recommended against using it, in order not to provide the government with excuses to use it against activists. Below are the terms identified and their definitions.

- Sexist and misogynistic: hate speech related to expressions that spread, incite, promote, or justify hatred based on sex.
- Degradation via irony and sarcasm: Although sarcasm is s part of freedom of speech, in some cases it can be weaponised to degrade people.
- Insult and bullying: name-calling, insults, teasing, intimidation, homophobic or racist remarks, or verbal abuse.
- Dehumanisation: depriving a person or group of positive human qualities.
- Demonisation: portraying someone as wicked and threatening.
- Stereotyping: relating to a widely held but fixed and oversimplified image or idea of a particular type of person or thing.
- Sexual harassment: behavior characterised by the making of unwelcome and inappropriate sexual remarks or physical advances.
- Physical appearance: negative labelling of an element of a person's appearance.
- Accusations: a charge or claim that someone has done something illegal or wrong.
- Information threat: The threat of hacking, sharing private information, blackmail or doxing.

After the content has been classified through the first two layers, annotators then go through a third layer of classification of the content, as illustrated in the mind-map mentioned previously, which is the “OGBV Basis”, in order to identify the basis of the OGBV and the causes of the hate speech against women on the Internet in Jordan. The annotations include discrimination on the basis of race, national origin, social and economic class, disability, religion, and gender identity/gender expression.

Furthermore, JOSA adopted the intensity scale shown below Figure 02, which has been used by other regional partners within the “Words Matter” project, as a reference for analysing the intensity of identified hate speech for the purpose of carrying out analysis, producing findings, and reporting. The intensity scale was initially measured based on the Social Science Research Council (SSRC) paper on "Classifying and Identifying the Intensity of Hate Speech [11].



Color	Title	Description	Examples
🟢	1 Disagreement	Rhetoric includes disagreeing at the idea/belief level. Responses include challenging claims, ideas, beliefs, or trying to change their view.	False, incorrect, wrong, challenge, persuade, change minds
🟡	2 Negative Actions	Rhetoric includes negative nonviolent actions associated with the group. Responses include nonviolent actions including metaphors.	Threatened, stole, outrageous act, poor treatment, alienate
🟠	3 Negative Character	Rhetoric includes nonviolent characterizations and insults. There are no responses for #3.	Stupid, thief, aggressor, fake, crazy
🟠	4 Demonizing and Dehumanizing	Rhetoric includes subhuman and superhuman characteristics. There are no responses for #4.	Rat, monkey, Nazi, demon, cancer, monster
🔴	5 Violence	Rhetoric includes infliction of physical harm or metaphoric/ aspirational physical harm or death. Responses include calls for literal violence or metaphoric/aspirational physical harm or death.	Punched, raped, starved, torturing, mugging
⚫	6 Death	Rhetoric includes literal killing by group. Responses include the literal death/elimination of a group.	Killed, annihilate, destroy

Figure 02: Classifying and Identifying the intensity of hate speech by the Social Science Research Council (SSRC)

The Training Phase

Following the completion of the first phase, JOSA proceeded with the second, which mainly focused on monitoring social media trends and hashtags, data collection, and data annotation.

The social media trends and hashtag monitoring was completed by JOSA's research, using the CrowdTangle tool, which is a tool provided by Meta that helps in following, analysing, and reporting on what is happening across social media, by tracking engagement

[11] Babak Bahador, "Classifying and Identifying the Intensity of Hate Speech", Social Science Research Council, 17 November 2020.

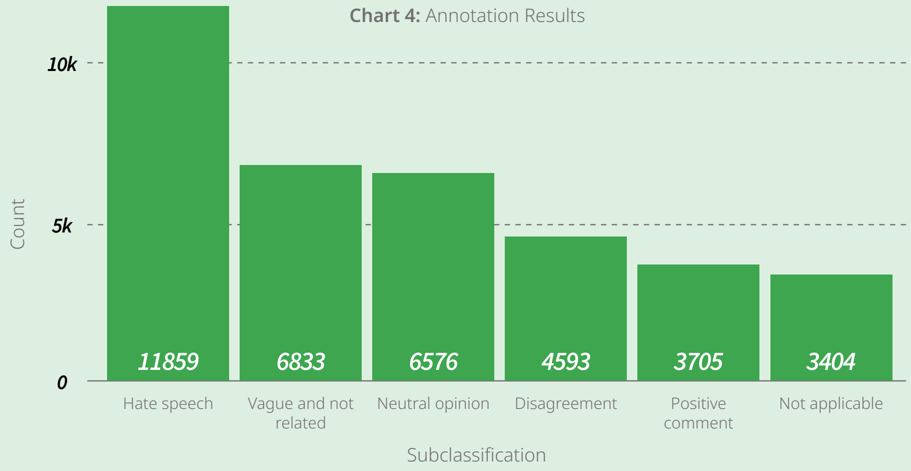
(interactions), which includes reactions (e.g., “Likes” on Facebook), comments, and shares.

Up until the first quarter of 2023, JOSA closely monitored over 20 per cent of the identified sample, constituting 260 Facebook posts, to train the Nuha model. JOSA intended to collect more content from Twitter, but was not able to secure access to the Twitter API, due to changes in Twitter policies (see the limitations section for more details). While collecting data, JOSA searched for and collected posts that were published over the past year (2022). The process of data collection was carried out in three steps requiring direct human intervention, with the support of online tools:

- Step 1: URL collection, using the CrowdTangle tool.
- Step 2: Content extraction, using the Export Comment tool; [12] which is used to scrape content from social media platforms (e.g., Facebook, Twitter, etc.).
- Step 3: Data cleaning, ensuring that all unnecessary data is deleted, and rearranging important data in a specific format.

Following the data collection, extraction, and cleaning, JOSA researchers, in collaboration with the TAMAM coalition, annotated approximately 37,000 Facebook comments as a first sample. The annotation was completed using the Label Studio annotation tool, relying on the mind-map and sub-classification identified in the previous phases (see the mind-map chart). The data sample monitored, annotated, and analysed only reflects the initial findings related to online gender-based hate speech in Jordan. The chart below (chart 04) reflects the first layer of annotation, as the percentage of hate speech for this sample was 32 per cent, with the remaining 68 per cent distributed across other categories – positive feedback, neutral opinion, disagreement, vague and not related, and not applicable.

Chart 4: Annotation Results



[12] Export Comments website

As mentioned previously, data was collected from women and women rights organizations in Jordan, and content was classified based on whether or not it contained online gender-based hate speech against women. Taking a closer look into the sub-classification of identified hate speech in the sample, in Chart 05, below, approximately 6,500 comments were annotated as insult and bullying, while approximately 3,200 were annotated as irony and sarcasm.

At the same time, death threats, harm threats, and information threats, which are considered highly dangerous in the intensity scale (figure 02), were the smallest group within the data sample, which we believe is related to social media content moderation.

Approximately 20 percent of the overall sample size and percentages only reflect the initial hate speech analysis. The Nuha team believes that these percentages could change during the next phase of data collection and can climb higher during major political and national events with heavy news coverage.

To better understand the process of annotation and what we mean by the specific sub-classifications of hate speech, below are some screenshots (Figures 3, 4, 5, and 6) of what was captured during the data monitoring, showing instances of insult and bullying, irony and sarcasm, and death threats.

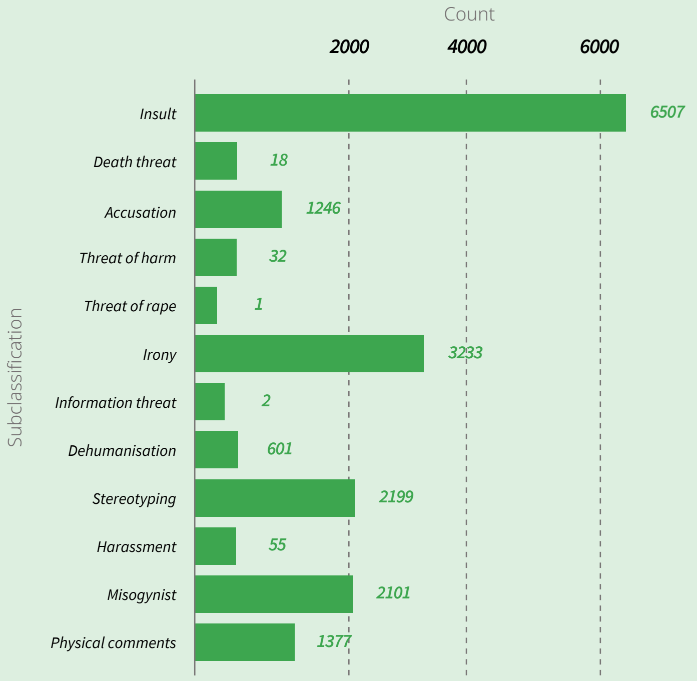
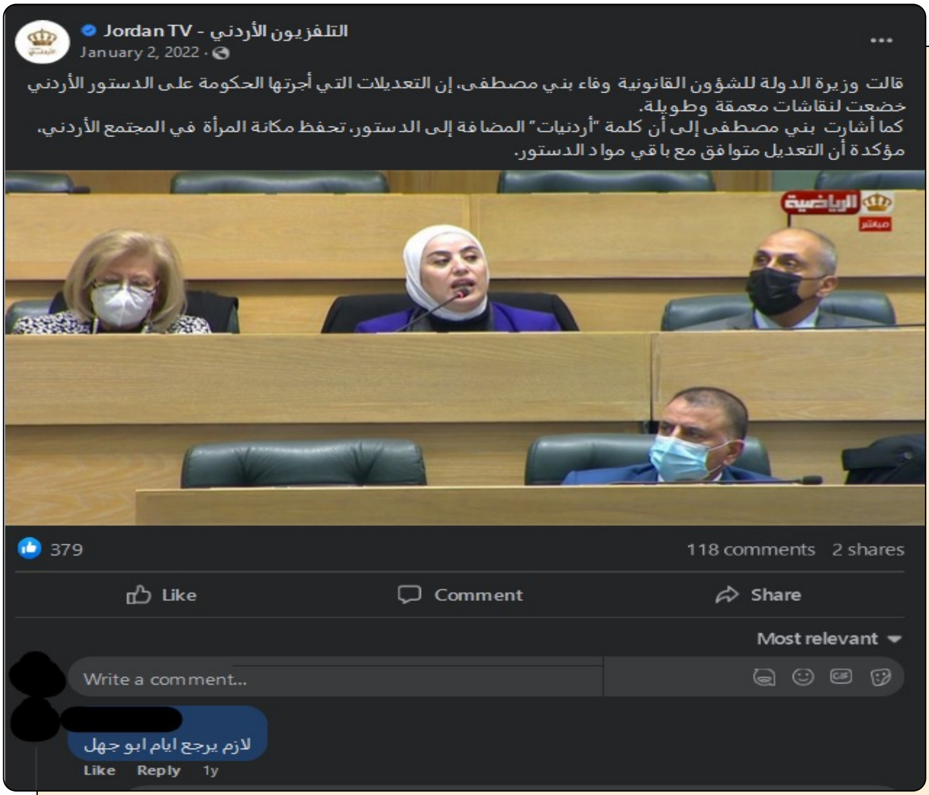


Chart 5: Hate speech sub-classification's analysis



Jordan TV - التلفزيون الأردني
January 2, 2022 ·

قالت وزيرة الدولة للشؤون القانونية وفاء بني مصطفى، إن التعديلات التي أجرتها الحكومة على الدستور الأردني
خضعت لنقاشات معمقة وطويلة.
كما أشارت بني مصطفى إلى أن كلمة "أردنيات" المضافة إلى الدستور، تحفظ مكانة المرأة في المجتمع الأردني،
مؤكدة أن التعديل متوافق مع باقي مواد الدستور.

[redacted]

[redacted]

لازم يرجع ايام ابو جهل

Figure 03: Example of a death threat

Date of post: 06/01/2022

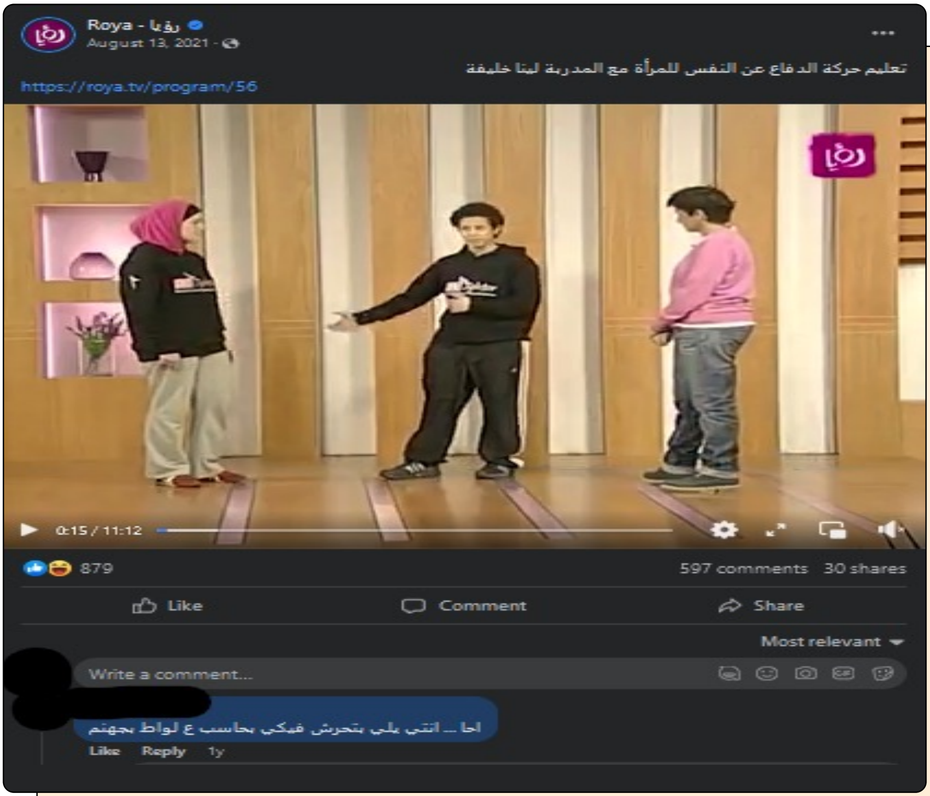
Time of comment: one year ago

Post text in English: "Minister of State for Legal Affairs Wafaa Bani Mustafa said that the amendments made by the government to the Jordanian Constitution were subject to in-depth and long discussions. Bani Mustafa also indicated that the words 'Jordanian women' added to the constitution preserve the status of women in Jordanian society, stressing that the amendment is compatible with the rest of the articles of the constitution"

Comment in English: "We must go back to the days of Abu Jahl"

Context: This comment was classified as "hate speech" and subclassified as a "death threat". The basis for this classification was "discrimination based on gender identity and expression", because the comment is referring to the pre-Islamic habit of female infanticide. Users who attack women on social media use metaphors to by-pass content-moderation algorithms.

Link to the post and comment: Here.



Roya - رؤيا
August 13, 2021

تعليم حركة الدفاع عن النفس للمرأة مع المدربة لينا خليفة

https://roya.tv/program/56

[redacted]

احا ... انتي يلي بتحرش فيكي بحاسب ع لواط بجهنم

Figure 04: Example of insult and bullying, and stereotyping

Date of post: 12/08/2021

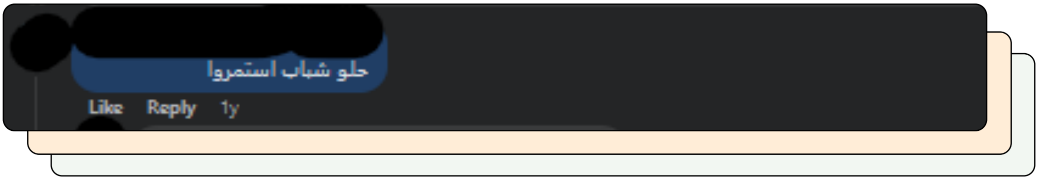
Time of comment: one year ago

Post text in English: "Teaching self-defense movements to women, with trainer Lina Khalifa"

Comment in English: "We must go back to the days of Abu Jahl"

Context: This comment was classified as "hate speech" and subclassified as "insult and bullying" and "stereotyping". The basis for this classification was "discrimination based on gender identity and expression", because the comment is insulting to the trainer, when the "trainer looks so masculine that she could almost be a man", inferring that, if a man harasses her, he would go to hell for homosexuality because it would be as if he harassed a man.

Link of the post and comment: Here.



[redacted]

حلو شباب استمروا

Like Reply 1y

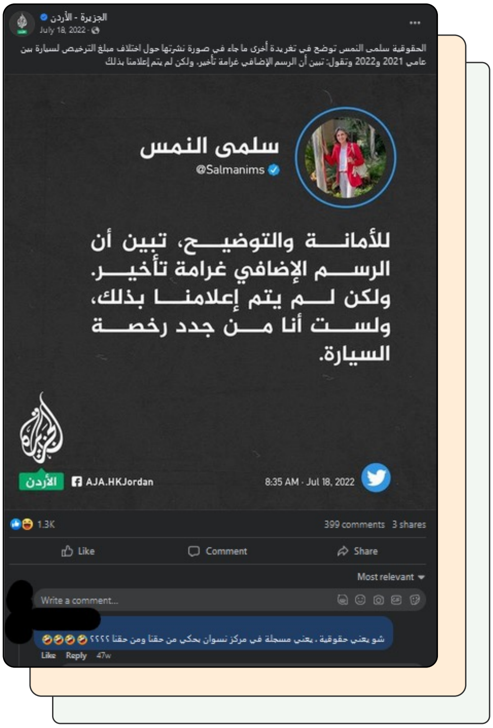
Figure 05: Example of irony and sarcasm, and misogynist or sexist discourse

This comment is on the same post as above.

Time of comment: one year ago

Comment in English: "Sweet guys, keep going"

Context: This comment was classified as “hate speech” and subclassified as “irony and sarcasm” and “misogynist or sexist”. The basis for this classification was “discrimination based on gender identity or expression”, because the commenter is referring to the women mentioned in the post as “men” because they know self-defense and they chose a non-conventionally “feminine masculine” way of expressing themselves as women, which makes them “men”, according to the person who posted this sarcastic and ironic comment.



الجزيرة - الأردن
July 18, 2022 ·

الحقوقية سلمى النمس توضح في تغريدة أخرى ما جاء في صورة نشرتها حول اختلاف مبلغ الترخيص لسيارة بين عامي 2021 و2022 وتقول: تبين أن الرسم الإضافي غرامة تأخير، ولكن لم يتم إعلامنا بذلك

سلمى النمس
@Salmanims

للأمانة والتوضيح، تبين أن الرسم الإضافي غرامة تأخير. ولكن لم يتم إعلامنا بذلك، ولست أنا من جدد رخصة السيارة.

8:35 AM · Jul 18, 2022

1.3K 399 comments 3 shares

[redacted]

شو يعني حقوقية . يعني مسجلة في مركز نسوان بحكي من حقنا ومن حقنا ؟؟؟؟ 🤣🤣🤣🤣

Like Reply 47w

Figure 06: Example of irony and sarcasm

Date of post: 16/07/2022

Time of comment: 47 weeks ago

Post text in English: "Human rights activist Salma Al-Nims explains in another tweet what was shown in a photo she published ab out the difference in the cost of a license for a car from 2021 to 2022, saying: 'It turned out that the additional fee is a delay fine, but we were not notified of that'."

Comment in English: “What does it mean to be a human rights activist? It means she is registered at the Women’s Centre. To say I am talking about this is our right and that’s our right????"

Context: This comment was classified as “hate speech” and subclassified as “irony and sarcasm”. The basis for this classification was “discrimination based on gender identity and expression", because the commentor is making fun of women rights activists, saying that they are not real human rights activists and making a mockery or her activism as a feminist.

Link of the post and comment: Here.

Deployment Phase

After the above-mentioned two phases, JOSA began by building the model and training it with the annotated data to detect hate speech. Up to the reporting period, JOSA had developed a model with a 72 per cent F1 score, and will continue working on increasing that score, aiming for 90 – 95 per cent. Building and training the model process was completed by dividing the annotated data into training and validation sets. In order for the model to learn the patterns and features of content containing hate speech against women, as well as content that does not, JOSA gave the model the training set. After the model had been trained on that data, JOSA began testing it, using the validation set of data, where it was given data that had not been trained on before to test its ability to detect hate speech.

However, during the first trial of developing and training the Nuha model, relying on the first sample of annotated data, we encountered a challenge in the imbalance between the number of samples classified as "hate speech" and those classified as "non-hate speech". The imbalance between the two categories is likely one of the main factors leading to the F1 score achieved, along with the sample size itself.

Considering all of the factors involved in developing the tool (the number of monitored accounts, the size of the sample, and the percentages of hate speech and non-hate speech content), a 72 per cent F1 score is a good result for the first trial. Nonetheless, JOSA is aiming to develop a tool with 90 – 95 per cent score. To further increase the F1 level, data collectors continued monitoring the identified accounts, initiatives, and hashtags, and increased the data sample to 260 posts, with more than 37,000 FB comments and tweet replies (Securing Twitter API approval remains a challenge to enhancing the data sample).

Along with increasing the sample, JOSA used data augmentation techniques to increase the F1 score of the model. The process of data augmentation enables us to artificially increase the size of the training data size by generating different versions of real datasets, i.e., utilising certain data samples to create additional but similar samples that could be used to train the Nuha model. In other words, data augmentation allows Nuha to learn from different versions of the same sample. Specifically, we utilised an open-source model called AraBERT (produced by the American University of Beirut) to substitute some words within the hate speech comments with semantically similar words. As an example of this data augmentation, if the original text was "I have no time", the augmented text carrying the same meaning, but in a different version, would be "I don't have time".

Monitoring Phase

Upon the completion Nuha's development and release, JOSA is planning an advocacy campaign that aims to spread the awareness of the model, the value it offers, and how to use it in different contexts. The campaign will target researchers, women's rights organisations, and others interested, through training sessions, roundtables, and webinars.

Key Findings:

Throughout the machine learning lifecycle of developing Nuha, JOSA's work resulted in several initial key findings. One of our main observations focused on the reasons that might be affecting types of hate speech on the internet, specifically Meta platforms. JOSA found a huge variance between Negative Action (yellow) and Negative Character (orange), verses Violence (red) and Death (black) within the annotated data sample (see Figure 02) – classifying and identifying the intensity of hate speech. This is likely for the following reasons:

- Increasingly effective content moderation by Meta.
- Violent content being deleted by owners after the posting.
- Existing laws and regulations; the Cyber-crime law in Jordan [13] might deter general internet users from online violence with such severe intensity.
- Meta's hate speech policy [14] does not consider the negative action, negative character, and demonising and dehumanising categories (see: Figure 02 – classifying and identifying the intensity of hate speech) as hate speech, but, rather, as freedom of expression.

The observations above were made based on the first annotated sample observation. As reflected in Chart 05 – "Hate speech sub-classification analysis", approximately 55 per cent of the content was annotated as insult and bullying, while 27 per cent was sub-classified as irony and sarcasm. On the other hand, death threats, threats of harm, threats of rape, and information threats, taken together, did not exceed 0.5 per cent. This does not imply, however, that women in Jordan are not subjected to death threats on the internet, but only reflects the initial threats JOSA monitored up until the first quarter of 2023. Regarding the hate speech narratives that were studied during the development of Nuha, the findings can be grouped within the following categories:

1. Violent criticisms of women based on their lifestyles (e.g., the clothes that they wear or the type of activities they are involved in, such as playing football).
2. Hate speech was directed against women who serve in jobs that are stereotypically considered male-only jobs in Jordanian culture.
3. Women who express their opinion on religious issues or discuss religious issues.

Most importantly, based on our analysis, a considerable number of women's posts and tweets on social media platforms that were subjected to these hostile comments were eventually deleted. Specifically, our examination of the collected data revealed that roughly **11 per cent of content published by women (September 2022 – March 2023) was no longer available online.** [16] It is unlikely that social media companies removed this content, as it does not

[13] https://www.jcca.org.jo/DataFiles/2017-10/2017/%D%82%9D%8A%7D%86%9D%88%9D%20%86%9D%8A%7D%84%9D%8AC%D%8B%1D8%A%7D%8A%6D%20%85%9D%8A%7D%84%9D%8A%7D%84%9D%83%9D%8AA%D%8B%1D%88%9D%86%9D8%9A%D%8A9.pdf

[14] Meta Transparency Center, "Hate Speech".

[15] Violent Criticism: The act of denouncing; public menace or accusation; the act of inveighing against, stigmatising or publicly arraigning.

violate their community standards. One possible explanation for this that the women, themselves, opted to remove the content in question, to avoid backlash and hate speech directed towards them. This finding is consistent with a report published by the Jordanian National Commission for Women, [17] which indicated that a significant number of women active in political and public spaces online in Jordan eventually withdrew from these spaces. Further research is needed to explore who took down this content, and why. JOSA is enhancing its documentation techniques to better understand this phenomenon.

Analysis and Discussion

For almost a decade now, women in Jordan in political spaces have commonly faced misogynist hate speech and online attacks, due to cultural gender roles that restrict them to the domestic sphere of taking care of their family, husband, children, and houses. For example, nine years ago, MP Hind Al-Fayez was subjected to sexist remarks by a colleague, who told her to “shut up” and “sit down, Hind!”, [18] which is emblematic of the recurring pattern of sexist hate speech and gender-based violence against women in Jordan. Such misogynist remarks are also consistent with the findings documented by Maharat’s Words Matter partner in Lebanon July in 2022, when MP Halima Kaakour demanded to be allowed to speak during a vote, but Speaker Nabih Berri refused and responded by saying, "Sit down and shut up!"

Recently, another MP, Mayada Shreim, was the target of aggressive comments on social media platforms after leading a "Sulh" meeting [19] (i.e., a reconciliation meeting to resolve conflicts between tribes), as she is the first woman to lead such a process in Jordan. The response of the Jordanian community to Shreim's involvement in the Sulh process was stereotypical and revealed gender biases, as the vast majority of Jordanians who reacted to the news item believed that only men should be involved in tribal reconciliations.

This incident was among the cases JOHA has observed over the past few months, indicating that women politicians and activists, including feminist activists, are increasingly being subjected to hate speech online. The dataset we analysed revealed that hate speech against women in politics in Jordan takes one of two forms; either women are told to “return to the kitchen, where they belong”, which is a literal phrase that is commonly used in Jordan, or they are accused of importing and implementing Western agendas that aim to destroy the Jordanian community.

[16] In the initial stages of the Nuha project, the primary objective of data collection was to facilitate the training of our AI model. During the period spanning from September 2022 to March 2023, a comprehensive compilation of social media URLs was obtained through the utilisation of web scraping techniques. These posts were subsequently incorporated into an Excel file for importation into the annotation tool employed (Label Studio). However, upon commencing the preparation of this report, a retrospective examination of the Excel sheet revealed that approximately 11 percent of the posts were absent, as they had not been successfully retrieved by the web scraper. Subsequent investigation confirmed that the corresponding URLs had become invalid, signifying either their deletion or a modification in privacy settings, rendering them no longer publicly accessible.

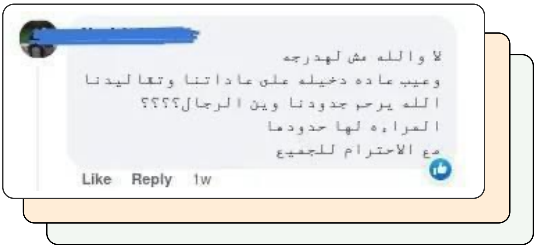
[17] The Jordanian National Commission for Women (2022), Violence Against Women in the Public and Political Spheres in Jordan.

[86] Al Arabiya News, “Jordan MP’s ‘Sit Down Hind’ Memes Go Viral”, 7 December 2014.

[19] Al Arabiya News, “Woman MP’s Leading of a Sulh Meeting Sparks Controversy in Jordan”, ??/??/2023.

Figure 07: MP Mayada Shreim during a "Sulh". (Photo credit: Ro'ya News)

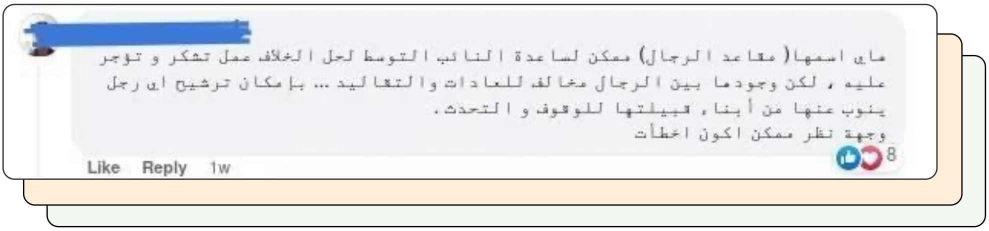
Our study of the data also revealed instances of highly violent criticisms that could be classified as defamation and hate speech against women based on their social practices, physical appearance, and social traditions and customs, such as their clothing or the type of activities they engage in. This issue is most prevalent in content related to Jordanian women in sports. For example, Ayah Majali, a Jordanian football player, was the target of derogatory comments for wearing shorts. A similar situation arose with the Jordanian national women's basketball team. This form of harmful content is often manifested in derogatory "male guardianship" statements about the players' male family members, insinuating that the women "do not have a single man in the family".



[redacted]

لا والله مش لهدرجه
وعيب عاده دخيله على عاداتنا وتقاليدنا
الله يرحم جدودنا وين الرجال؟؟؟؟
المراءه لها حدودها
مع الاحترام للجميع

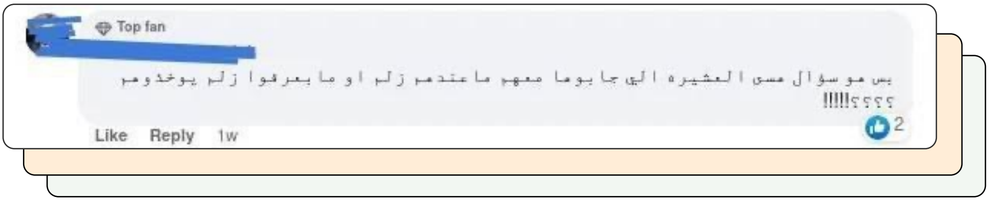
Translation: "I swear to God, No, This is too much. This is shameful, it's not within our norms and traditions. Mercy on the souls of our grandparents, where are the men???? Women have limits, With all respect to all". The commenter's main message was that women should not take roles of men and that they have lines they cannot cross, and suggests that there are no men left to fulfill those social roles.



[redacted]

هاي اسمها (مقاعد الرجال) ممكن لساعدة النائب التوسط لحل الخلاف عمل تشكر و تؤجر
عليه ، لكن وجودها بين الرجال مخالف للعادات والتقاليد ... بإمكان ترشيح اي رجل
ينوب عنها من أبناء قبيلتها للوقوف و التحدث .
وجهة نظر ممكن اكون اخطأت

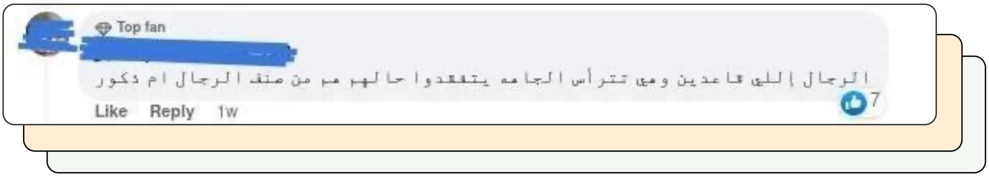
Translation: "These sessions are only for men, She [referring to MP Mayada] could at best contribute to resolve conflicts at work, but her being in a Sulh session between all those men goes against our norms and traditions.... She could have nominated a man from her tribe to speak on her behalf... That's my point of view, could be right or could be wrong as a result of norms and traditions.



Top fan
[redacted]

بس هو سوال هسى العشيره الي جابوها معهم ماعندهم زلم او ما بعرفوا زلم يوخذوهم
؟؟؟؟!!!!!

Translation: "Just a question in my mind, now the tribe that asked her to lead this session don't have men or at least allies from other tribes?" The commenter is questioning the tribe's decision to ask a women to lead the Sulh, as if there were no men to do so, or they do not know of any qualified men.



Top fan
[redacted]

الرجال إللي قاعدين وهي تترأس الجاهه يتفقدوا حالهم هم من صنف الرجال ام ذكور

Translation: "Men who are sitting there while she's leading the Sulh should double check whether or not they are men". The commenter is questioning men's masculinity in a sarcastic way in a situation where a woman took the lead.

Figure 08: People's comments on the news items about the "Sulh" session lead by MP Mayada Shreim

JOHA has observed a significant amount of hate speech directed towards Jordanian women who work in professions that are traditionally considered the domain of men. Rashan, an unemployed engineer who started working as a bus driver, is one such example. The comments directed at Rashan were particularly virulent, to the extent that our team felt compelled to warn data annotators about the graphic and violent nature of the comments before they starts annotating them.

Research Limitations

Twitter API Access Difficulty:

- Major challenge in Nuha tool development.
- Unable to get Twitter API approval.
- Used alternative tool "Export Comments."

Imbalance in Hate Speech Content:

- More non-hate speech content in the sample.
- Resulted in higher false positives and false negatives.
- Addressed using data augmentation and new content collection.

Lack of Prior Research:

- National-level gender and AI research absent in Jordan.
- Started with limited existing research.
- Difficulty in defining hate speech-related concepts.

Conclusion

Our preliminary analysis of the annotated data to train the Nuha AI model revealed that hate speech and harmful content against women in Jordan are prevalent in various contexts. Our study of the hate speech narratives that were identified during the development of Nuha reveals several categories, including attacks against women in politics and feminist activists, aggressive criticism of women based on their social practices and behavior, such as the clothes they wear or the type of activities they are involved in, hate speech directed against women who serve in jobs and/or roles that are considered male-only in Jordanian culture, and women who express their opinions on religious issues.

It is important to note that the analysis is still ongoing, and JOSA is continuing to study the data for a more developed analysis. JOSA, is still working on the data to form a clearer picture of hate speech against women in Jordan in digital spaces, alongside with our work on getting Nuha ready for its first release.

Recommendations

At this point, JOSA recommends the following:

- Academic institutions should prioritise the study of the intersections between gender and technology on a national level in Jordan. The absence of research papers addressing these issues is a significant gap in the academic discourse, and must be addressed in order to ensure safe and flourishing digital spaces, as well as the promotion of democracy and protection of human rights in the future.

- Decision-makers in Jordan, (such as ministries, legislators, etc.) must also commit to further studying this issue on a national level. We urge them to work closely with organisations focused on women's rights to develop frameworks that protect and empower women, enabling them to participate more effectively in digital public spaces without fear of bullying, cyber-violence, and hate speech.
- Provide front-line workers in Jordan with the skills and knowledge needed for them to be able to better deal with cases of OGBV.
- Social media platforms must also take responsibility and develop more gender-sensitive mechanisms to deal with online hate speech. While some platforms have made progress in this regard, there is still much to be done. It is essential to develop more effective mechanisms to respond to reports, to ensure the effective participation of women in the political and public digital spaces.
- Social media platforms should work closely with local civil society organisations to develop rapid-response helplines to support in taking down extreme content against women.
- Civil society in Jordan should work closely with the government to raise awareness about reporting hate crimes against women, and to provide support to targeted women online.
- Civil society should explore collaboration with religious institutes to mitigate the usage of religious speech to demonise and dehumanise women online.

3. Al Hayat Center (RASED)

Hate Speech and Disinformation on Social Media-Gender Lens

Context

The context in Jordan:

According to a 2018 report by the Jordanian National Commission for Women (JNCW), women in Jordan face significant challenges in terms of gender-based discrimination and violence, both online and offline. Following a previous report by Al Hayat Center – RASED, this report highlights the prevalence of gendered hate speech and gendered disinformation on social media platforms in the country, and the negative impact that it has on women's rights and empowerment.

This report seeks to address the prevalence and nature of gender-based hate speech in online discussions related to women in leadership positions in Jordan. Specifically, the study focuses on analysing comments made on Facebook posts related to statements made by Wafaa Bani Mustafa in July 2022, when she was Jordan’s Minister of legal affairs. The goal of the study is to shed light on the extent to which gender-based hate speech is used to discredit and undermine women in leadership positions in the country, and to provide recommendations for addressing this issue.

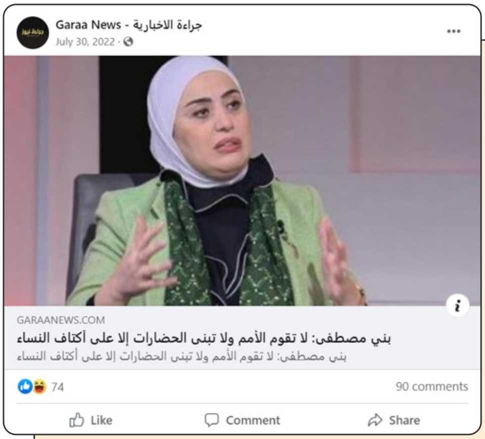
Who is Wafaa Bani Mustafa?

Wafaa Bani Mustafa – Minister of Social Development.

Figure 01: HE. Ms. Wafaa Bani Mustafa

Wafaa Bani Mustafa is a prominent Jordanian politician and lawyer who has held several crucial positions throughout her career. She has served as a member of the 16th, 17th, and 18th Jordanian Parliament, where she played a pivotal role in the country's legislative process.

Politically, Bani Mustafa has been known to follow a national political line and a moderate Islamic line, and she is from a conservative social environment outside of Amman, the capital. Her opinions expressed in Parliament have been in support of conservative social traditions and the religious framework.



Garaa News - جراءة الاخبارية

July 30, 2022

GARAANEWS.COM

بني مصطفى: لا تقوم الأمم ولا تبنى الحضارات إلا على أكتاف النساء

بني مصطفى: لا تقوم الأمم ولا تبنى الحضارات إلا على أكتاف النساء

74

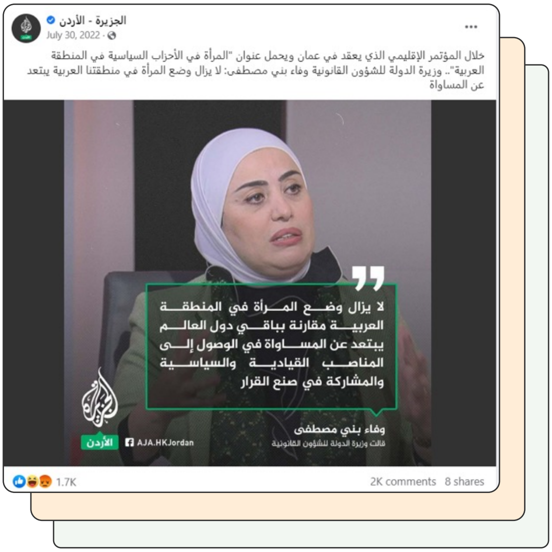
90 comments

Figure 02: Screenshot from "Garaa News Com's Post on Facebook".

The context of Mustafa's statements:

The statements made by Bani Mustafa, who was the Minister of State for Legal Affairs and the Chairperson of the Ministerial Committee for the

Empowerment of Women at the time, came during the regional conference “Women in Political Parties in the Arab Region,” held in Amman from 29-30 July 2022 [20], where she said, “Civilisations are not built except on the shoulders of women.” She added that “the status of women in the Arab region, compared to the rest of the world, is still far from equality in access to leadership and political positions and participation in decision-making.”



الجزيرة - الأردن
July 30, 2022 ·

خلال المؤتمر الإقليمي الذي يعقد في عمان ويحمل عنوان "المرأة في الأحزاب السياسية في المنطقة العربية".. وزيرة الدولة للشؤون القانونية وفاء بني مصطفى: لا يزال وضع المرأة في منطقتنا العربية يبتعد عن المساواة

لا يزال وضع المرأة في المنطقة العربية مقارنة بباقي دول العالم يبتعد عن المساواة في الوصول إلى المناصب القيادية والسياسية والمشاركة في صنع القرار

وفاء بني مصطفى
قالت وزيرة الدولة للشؤون القانونية

Figure 03: Screenshot from "AJA HKJordan (AlJazeera) Post on Facebook".

Bani Mustafa made a previous statement on July 17, 2022, during the closing symposium held by the Arab Women’s Center for Research and Training and the Jordanian National Committee for Women’s Affairs, entitled “Women Business Owners Economically and Commercially on an Equality with Men,” within a project to empower women entrepreneurs in the Middle East and North Africa. She said, “We must move forward in achieving more legislative justice for women, especially in legislation related to their political and economic participation, which has become an urgent necessity to increase family income, move the economy and increase the GDP.”

These are the words that were taken out of their general context and quoted as random statements that might cause controversy, opening the door for attacking Bani Mustafa in the aforementioned posts.

[20] Maria Weldali, “Conference Hears Challenges on Path to Women’s Political Participation”, The Jordan Times, 18 August 2023.

Methodology:

Collecting and Categorising Data on Gender-Based Hate Speech:

This report on gender-based hate speech was prepared through desk research. To collect data, specific keywords such as (وفاء بني مصطفى، المرأة، مساواة) translated as "Wafaa Bani Mustafa," "Women," and “Equality”, were used to search for relevant posts on Facebook, using CrowdTangle. After collecting the posts' URLs, the data were cleaned, filtered, and shortened to seven main posts out of the 33 that were exported by CrowdTangle, with a total number of 3,994 comments.

Out of the 3,994 comments on the 7 main posts, the research team was able to manually collect and categorise 3,126 comments, with the remainder being hidden or deleted during the data collection process.

These comments were categorised based on the type of gender-based hate speech being used. This categorisation process involved analysing the language used in each comment to identify instances of sexist, misogynistic, or other forms of gender-based hate speech. The Social Media Monitoring (SMM) team also took into account the context in which the comments were made, and the intended target of the hate speech.

To strengthen these findings, the team used the Export Comments tool to collect more than 3,500 comments, including nested comments, which are those comments made in response to other comments within a thread or discussion. These nested comments provide additional context and perspective.

To analyse and categorise the collected comments, the data analyst used Open AI Label Studio, which is a tool that facilitates the annotation and labelling of data for machine learning purposes. It allows users to define labels or categories, and then assign those labels to different data points.

After categorising the comments, it was found that over 30 per cent contained hate speech towards Bani Mustapha. The analysis of the comments and the use of Open AI Label Studio helped identify key findings related to the text of the comments and the GIFs and pictures used in them.

Overall, this methodology allowed the SMM team to collect and categorise data on gender-based hate speech and gain insights into the types of hate speech that are prevalent on social media platforms in Jordan, specifically, in this case study, towards “HE. Ms. Wafaa’ Bani Mustafa”.

Data sample:

For this report, the scope was limited to seven Facebook posts related to statements made by the Minister of Social Development (the Minister of State for Legal Affairs at that time) Wafaa Bani Mustafa, in July 2022. A total of 3,126 comments were analysed by the SMM team, and the percentage of comments containing gendered hate speech was found to be 33.4 per cent.

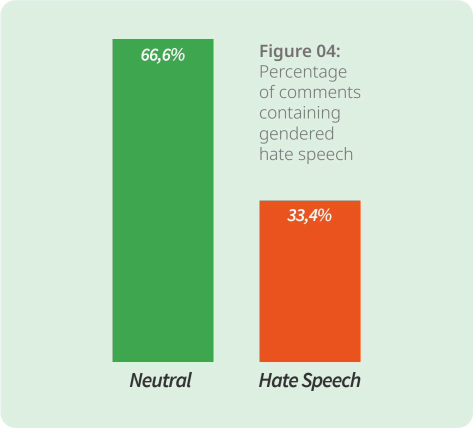
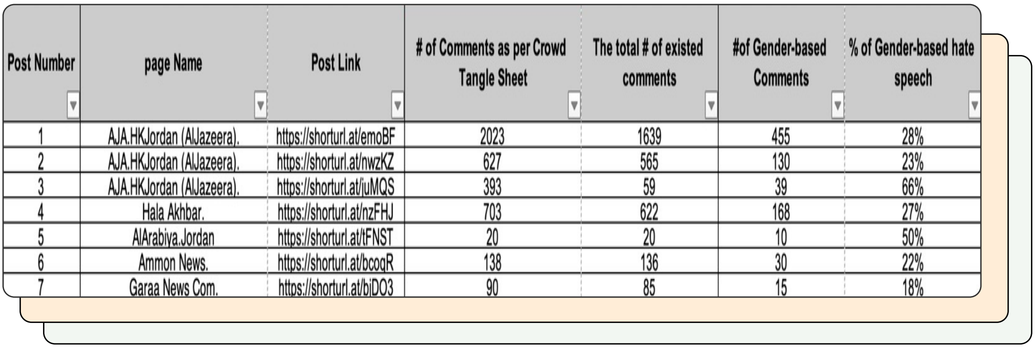


Figure 04: Percentage of comments containing gendered hate speech

The seven main posts analysed in this report on gender-based hate speech were published by several media entities in Jordan, including GaraaNewsCom, Ammonnews, AJA.HK Jordan, AlArabiya. Jordan, and HalaAkhbar.

The reach of these media entities is very wide, and they play an important role in shaping public opinion in Jordan, used as sources of information by much of the public.



Post Number	page Name	Post Link	# of Comments as per Crowd Tangle Sheet	The total # of existed comments	#of Gender-based Comments	% of Gender-based hate speech
1	AJA.HKJordan (AlJazeera).	https://shorturl.at/emoBF	2023	1639	455	28%
2	AJA.HKJordan (AlJazeera).	https://shorturl.at/nwzKZ	627	565	130	23%
3	AJA.HKJordan (AlJazeera).	https://shorturl.at/juMQS	393	59	39	66%
4	Hala Akhbar.	https://shorturl.at/nzFHJ	703	622	168	27%
5	AlArabiya.Jordan	https://shorturl.at/tFNST	20	20	10	50%
6	Ammon News.	https://shorturl.at/bcoqR	138	136	30	22%
7	Garaa News Com.	https://shorturl.at/bjDO3	90	85	15	18%

Data Analysis and Classification:

To analyse the data, the research team used a combination of qualitative and quantitative methods. Discourse analysis was conducted to identify patterns and themes in the comments, with a focus on identifying hate speech related to gender. The analysis was also guided by a gender analysis framework, which helped to highlight the ways in which gender intersects with other factors, such as ethnicity, religion, and socioeconomic status, in shaping experiences of hate speech.

In addition to qualitative analysis, the research team processed the data quantitatively to gain further insights. This involved categorising the types of hate speech, and the use of derogatory and offensive language, including slurs, insults, and threats, intended to denigrate, intimidate, or marginalise individuals based on gender. The team also calculated the frequency of use of each method and tool of gender-based hate speech across the 3,126 comments collected.

The research team also focused on ethical considerations, such as maintaining the anonymity of the commenters and ensuring that the study does not contribute to the further spread of hate speech and disinformation.

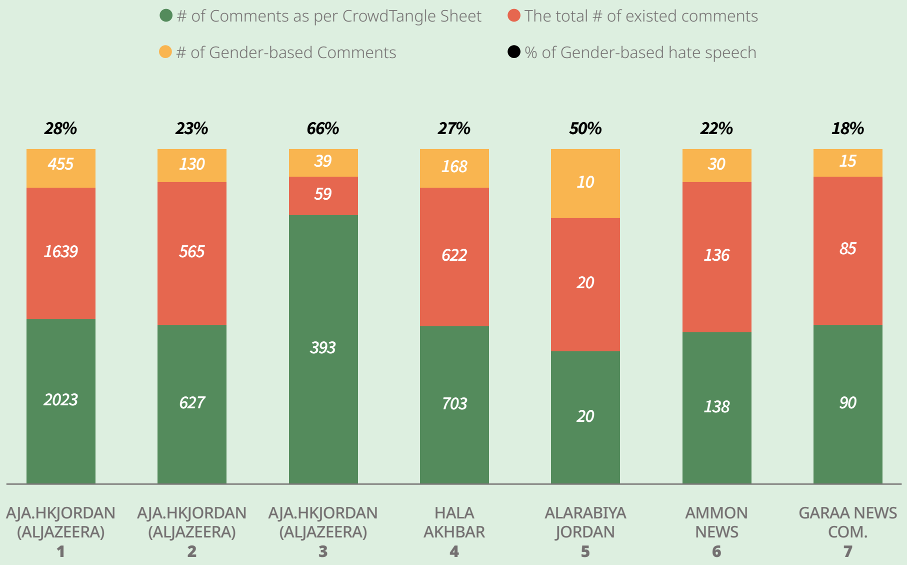


Figure 05: Data General Classification

Narratives in the analysis of the interaction with the statements of Minister Wafaa Bani Mustafa:

33.4 per cent of the 3,126 comments analyzed and found to contain Hate Speech.

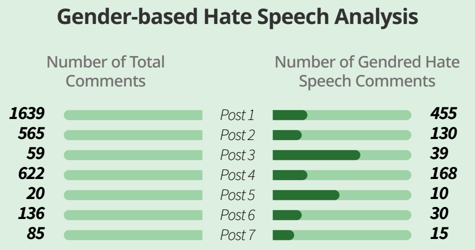


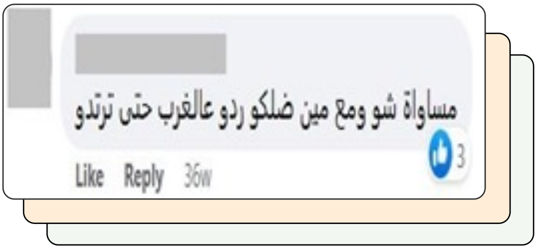
Figure 06: Gender-based Hate Speech Analysis

One significant observation was that many comments containing hate speech revolved around the belief that women should be confined to the kitchen and that, if a woman wears a hijab, she cannot express opinions about equality between women and men.

Interestingly, the analysis also revealed that many of the comments containing hate speech were written by women, indicating that the issue of gender-based hate speech towards women is not limited to men alone.

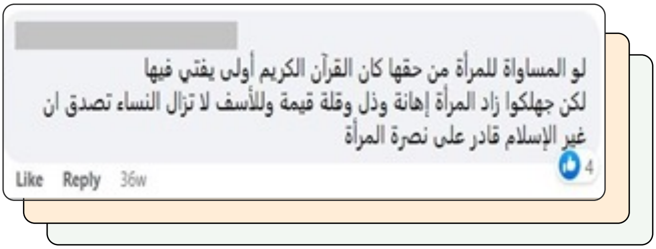
it was also observed that some male respondents exhibited a sarcastic tone toward women advocating for gender equality. These respondents argued that gender equality is not a pressing issue in Jordanian society, as men also suffer from low economic conditions and limited participation in decision-making positions. The following are some examples of these comments:

1. **The use of disapproving questions about the concept of equality, rejecting the idea of equality between men and women, and emphasising the idea of justice as stated in the Islamic religion; examples:**


[redacted]
[redacted]

مساواة شو ومع مين ضلكو ردو عالغرب حتى ترتدو

Translation: "Equality with whom? Keep listening to the west until you commit apostasy".

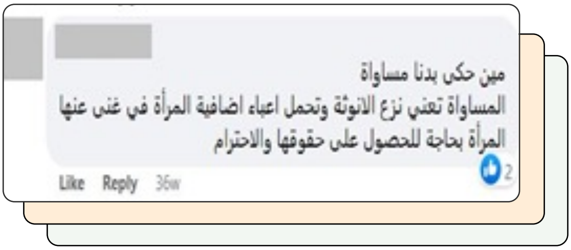

[redacted]

لو المساواة للمرأة من حقها كان القرآن الكريم أولى يفتي فيها
لكن جهلكوا زاد المرأة إهانة وذل وقلة قيمة وللأسف لا تزال النساء تصدق ان
غير الإسلام قادر على نصرة المرأة

Translation: "If women had the right to equality, the Quran would have given this to them already, but your ignorance contributed to humiliating women and degrading their value and, unfortunately, women still believe a religion other than Islam can safeguard and protect women".

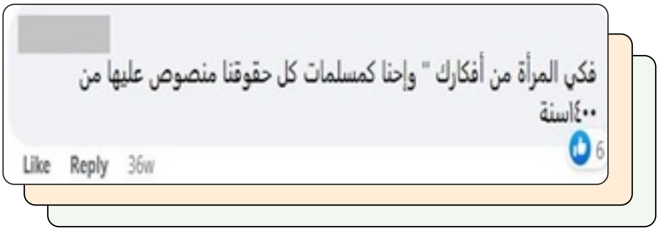
2. **Looking though a gender lens, we found that some users repeatedly commented negatively about the concept of "equality" between men and women, based on religious beliefs and traditional images of femininity; examples:**

   The analysis revealed that many comments containing hate speech were written by women themselves, indicating that the issue of gender-based hate speech towards women is not limited to men alone. The following screenshots are of comments written by women.



مين حكى بدنا مساواة
المساواة تعني نزع الانوثة وتحمل اعباء اضافية المرأة في غنى عنها
المرأة بحاجة للحصول على حقوقها والاحترام

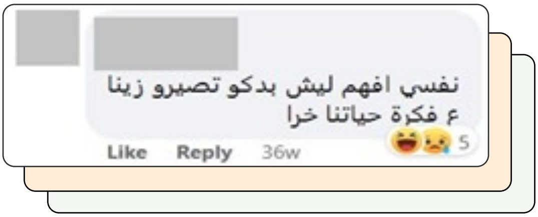
Translation: "Who said we want equality? Equality means taking the femininity of women away from them and giving them responsibilities that they are better off without. A woman needs to have her rights and respect."



فكي المرأة من أفكارك " واحنا كمسلمات كل حقوقنا منصوص عليها من
١٤٠٠سنة

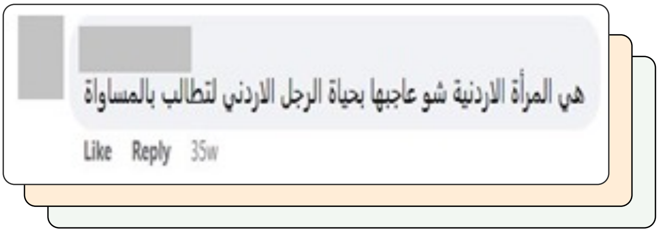
Translation: "Don't include women in your values, and we as Muslim women have had our rights established 1,400 years ago", meaning with the birth of Islam.

3. **it was observed that some male respondents exhibited a sarcastic tone toward women advocating for gender equality. These respondents argued that gender equality is not a pressing issue in Jordanian society, as men also suffer from low economic conditions and limited participation in decision-making positions; examples:**



نفسي افهم ليش بدكو تصيرو زينا
ع فكرة حياتنا خرا

Translation: "I want to understand why you [Jordanian women] want to be equal to us [Jordanian men]? By the way, our life is shitty."



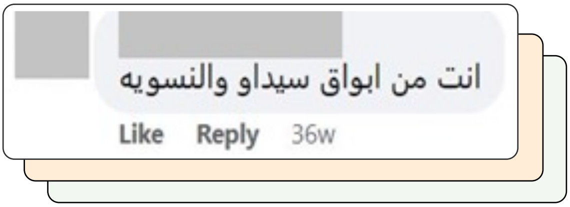
[redacted]

هي المرأة الاردنية شو عاجبها بحياة الرجل الاردني لتطالب بالمساواة

Like Reply 35w

Translation: "What is tempting about the life of the Jordanian man that makes Jordanian women demand for equality with him?"

4. Some respondents claimed that Bani Mustafa was implementing the agenda and agreements of foreign countries and organisations, particularly the Convention on the Elimination of All Forms of Discrimination Against Women (CEDAW)5, which they believe conflicts with the traditions and morals of Jordanian society. They expressed concern that this foreign influence is eroding Jordan's culture and values; examples:

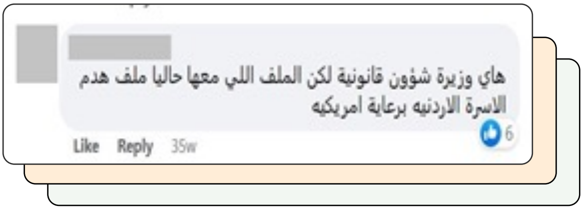


[redacted]

انت من ابواق سيداو والنسويه

Like Reply 36w

Translation: "You are one of the mouthpieces promoting CEDAW and feminism."



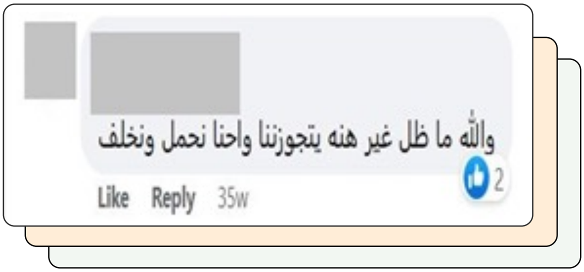
[redacted]

هاي وزيرة شؤون قانونية لكن الملف اللي معها حاليا ملف هدم
الاسرة الاردنيه برعاية امريكيه

Like Reply 35w 6

Translation: "This is the Minister of Legal Affairs, but what she is working on will lead to destroying the Jordanian family system, and with American sponsorship."

5. Some respondents expressed the view that women who call for gender equality should also take on physically demanding jobs, such as construction and loading and unloading work, and bear the same financial responsibilities as men, such as paying dowries and marriage costs; examples:

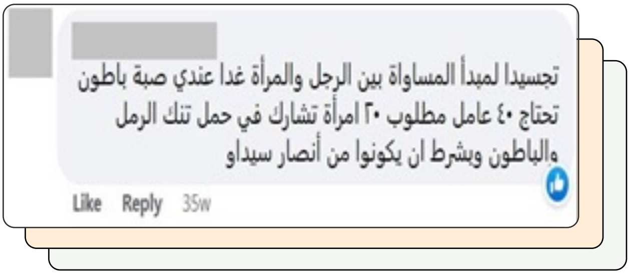


[redacted]

والله ما ظل غير هنه يتجوزنا واحنا نحمل ونخلف

Like Reply 35w 2

Translation: "The only thing that is left now is for women to marry us and we [men] to become the ones to get pregnant and give birth."

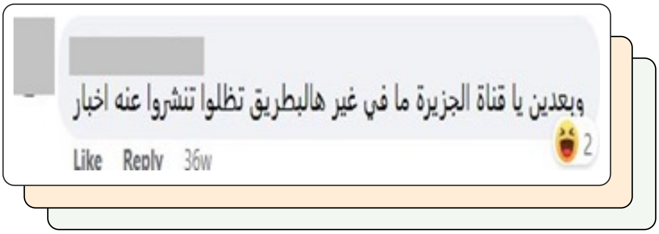


[redacted]

تجسيدا لمبدأ المساواة بين الرجل والمرأة غدا عندي صبة باطون
تحتاج ٤٠ عامل مطلوب ٢٠ امرأة تشارك في حمل تنك الرمل
والباطون وبشرط ان يكونوا من أنصار سيداو

Translation: "To embody the principle of equality between men and women, I have construction work in my house tomorrow that will require the efforts of 40 workers, and I'm looking for 20 women to participate in the construction and loading work, and with the condition that these women are supporters of CEDAW."

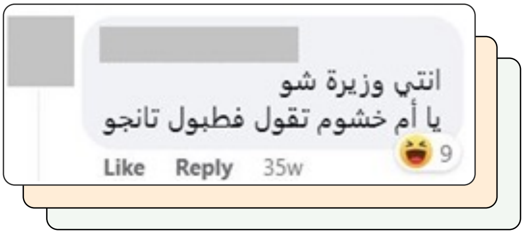
6. Some commenters resorted to using hate speech to criticise Bani Mustafa's physical appearance, employing derogatory and dehumanising language that compared her to animals; examples:



[redacted]

وبعدين يا قناة الجزيرة ما في غير هالبطريق تظلوا تنشروا عنه اخبار

Translation: "Why does Al Jazeera keeps covering news updates about this 'penguin'?"

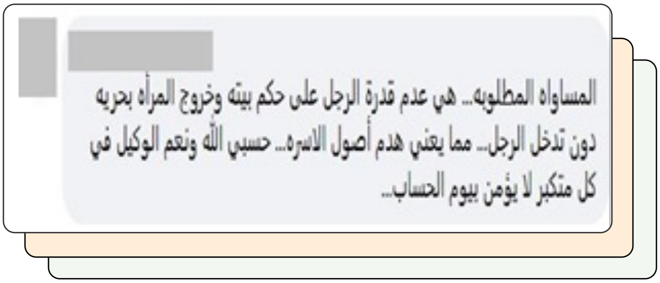


[redacted]

انتي وزيرة شو
يا أم خشوم تقول فطبول تانجو

Translation: "You are a minister of what – a women with big nose like you that looks like a tango football?"

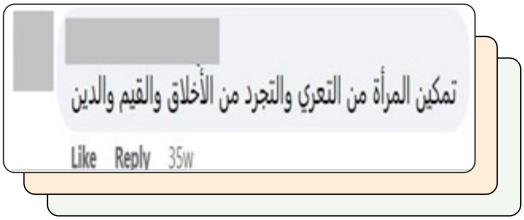
7. Some respondents expressed concerns that promoting gender equality and women's empowerment could lead to negative consequences, such as the disintegration of the family, the corruption of children, and deviation from societal customs and religious principles; examples:



[redacted] [redacted]

المساواه المطلوبه.. هي عدم قدرة الرجل على حكم بيته وخروج المرأه بحريه
دون تدخل الرجل.. مما يعني هدم أصول الاسره.. حسبي الله ونعم الوكيل في
كل متكبر لا يؤمن بيوم الحساب..

Translation: "The desired equality means that the man is unable to control his household, and women would have the freedom to go out without the man's approval, which means the destruction of the foundations of the family. He supplicated against those who do not believe in the Day of Judgement" (meaning that those who believe in and approve of gender equality will be punished on Judgement Day).

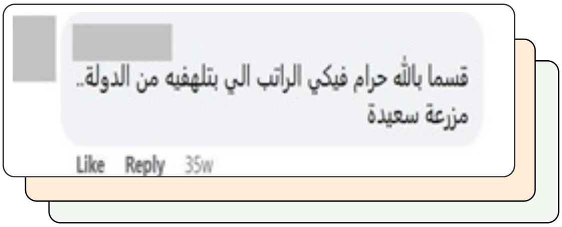


[redacted] [redacted]

تمكين المرأة من التعري والتجرد من الأخلاق والقيم والدين

Translation: "What is meant by empowering women is to strip them naked of their clothes and strip them of morals, values, and religion."

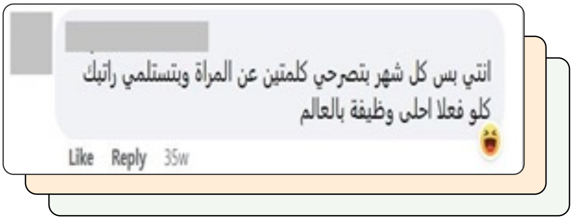
8. Some respondents criticised Bani Mustafa's statements about women's rights and her calls for gender equality, questioning the effectiveness of her work beyond making public statements. They alleged that her work is limited to giving speeches and making statements, and that she is overpaid for this; examples:



[redacted] [redacted]

قسما بالله حرام فيكي الراتب الي بتلهفيه من الدولة..
مزرعة سعيدة

Translation: "You don't deserve the salary paid by the government; this country is a happy farm" (a reference to the lack of accountability and prevalent corruption).



[redacted] [redacted]

انتي بس كل شهر بتصرحي كلمتين عن المراة وبتستلمي راتبك
كلو فعلا احلى وظيفة بالعالم
😆

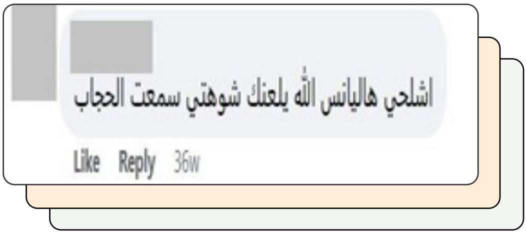
Translation: "Your job is to make statements about women's rights and receive a salary for it; indeed, this is really the best job in the world."

Discourse tools and methods:

During the analysis of the comments, the research team found many instances of hate speech in various forms such as "bullying, sexual harassment, defamation, insulting, and cursing" based on gender. The methods of discourse used in these interactions were as follows:

1. Negative use of religious discourse:

Critics of Bani Mustafa's advocacy of gender equality stated that she dares wearing a hijab, while her views are considered contrary to the principles of the Islamic religion and the customs and traditions of Jordanian society, which are mostly derived from Islamic strictures; examples:

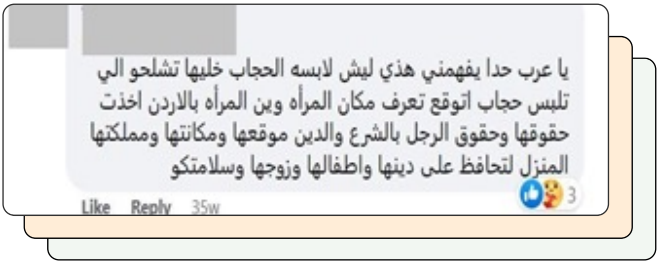


[redacted] [redacted]

اشلحي هاليانس الله يلعنك شوهتي سمعت الحجاب

Like Reply 36w

Translation: "May Allah curse you. You have harmed the reputation of the Hijab; take it off."

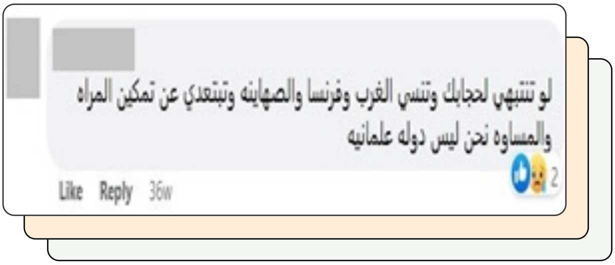


[redacted] [redacted]

يا عرب حدا يفهمني هذي ليش لابسه الحجاب خليها تشلحو الي
تلبس حجاب اتوقع تعرف مكان المراه وين المرأه بالاردن اخذت
حقوقها وحقوق الرجل بالشرع والدين موقعها ومكانتها ومملكتها
المنزل لتحافظ على دينها واطفالها وزوجها وسلامتكو

Like Reply 35w 3

Translation: "Why does this lady wear the Hijab? Women in Jordan have all of their rights, which she was given through religion. The place for a woman is home, where she keeps her religion and her children."

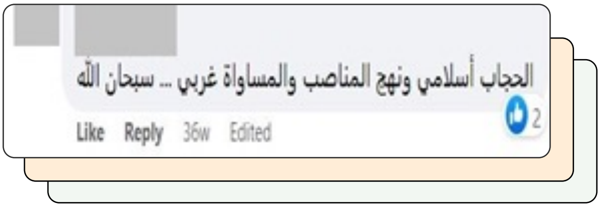


[redacted] [redacted]

لو تنتبهي لحجابك وتنسي الغرب وفرنسا والصهاينه وتبتعدي عن تمكين المراه
والمساوه نحن ليس دوله علمانيه

Like Reply 36w 2

Translation: "It is better for you to pay attention to your hijab and ignore the West, France, and Zionists, and distance yourself from empowering women and equality. We are not a secular state."



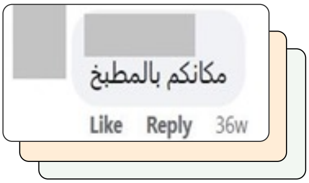
[redacted]

الحجاب أسلامي ونهج المناصب والمساواة غربي ... سبحان الله

Translation: "This woman wears the Islamic hijab and follows the European Western approach to having senior positions and gender equality. Subhan Allah (Glory be to God)!"

2. Gender-related stereotypes:

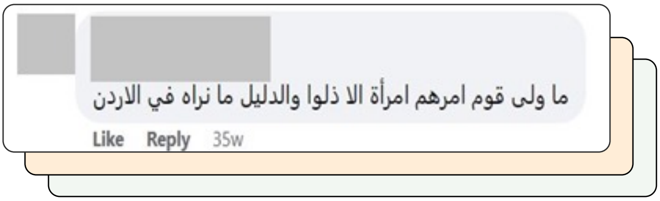
Bani Mustafa and women advocating for gender equality were subject to criticism based on stereotypes regarding the role of women in society; examples:



[redacted]

مكانكم بالمطبخ

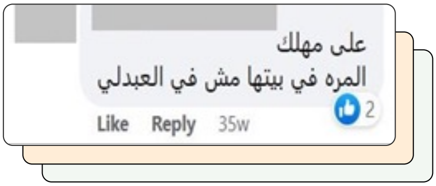
Translation: ""Women belong to the kitchen."



[redacted]

ما ولى قوم امرهم امرأة الا ذلوا والدليل ما نراه في الاردن

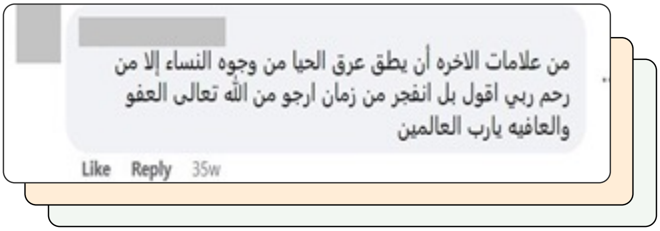
Translation: "Granting positions to women brings humiliation and disgrace to the people who do so, and the evidence is what is happening today in Jordan."



[redacted]

على مهلك
المره في بيتها مش في العبدلي

Translation: "Calm down; women should stay in the home, not in Abdali [referring to the headquarters of the Jordanian Parliament]."



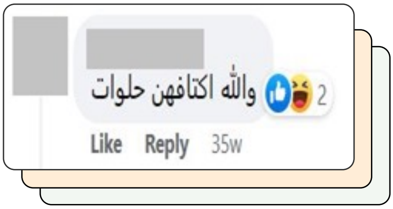
[redacted]

من علامات الاخره أن يطق عرق الحيا من وجوه النساء إلا من
رحم ربي اقول بل انفجر من زمان ارجو من الله تعالى العفو
والعافيه يارب العالمين

Translation: "One of the signs of Judgement Day is that women become immodest and abandon modesty. I beseech Allah for His mercy and forgiveness."

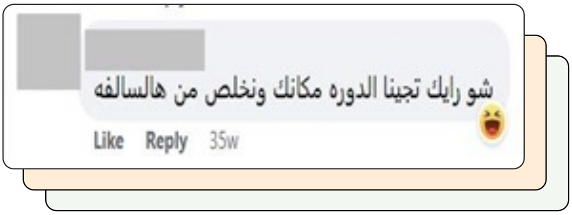
3. **The use of gender-based violence.**

A. **Sexual harassment:** The SMM team found many comments containing sexual insinuations, whether towards Bani Mustafa or women advocating for gender equality; examples:


[redacted]

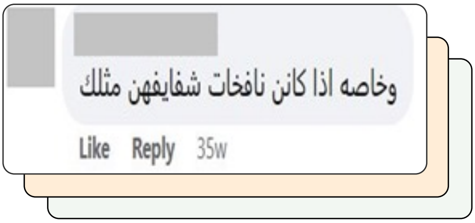
والله اكتافهن حلوات

Translation: "Ladies' shoulders are beautiful and seductive."


[redacted]

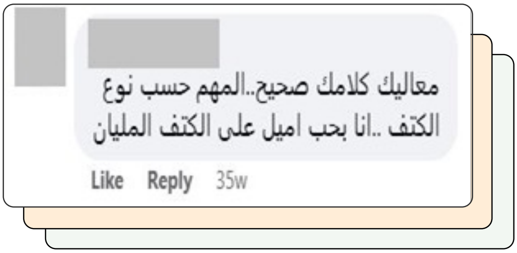
شو رايك تجينا الدوره مكانك ونخلص من هالسالفه

Translation: "I propose that the menstrual cycle be included as part of male biology, thereby ceasing discussions related to equality and women's rights. What do you think about this?"


[redacted]

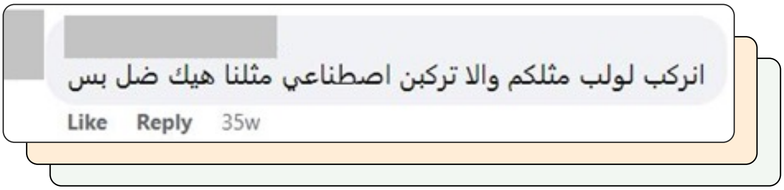
وخاصه اذا كانن نافخات شفايفهن مثلك

Translation: "Especially if the women have had cosmetic procedures for their lips."


[redacted]

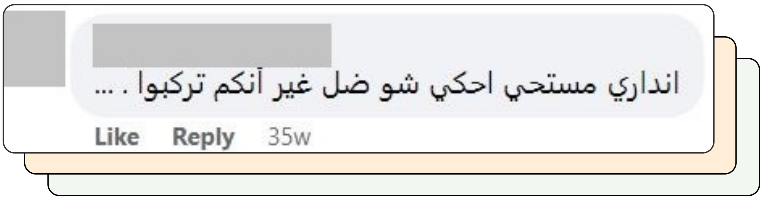
معاليك كلامك صحيح..المهم حسب نوع
الكتف ..انا بحب اميل على الكتف المليان

Translation: "You are correct, Your Excellency, but the type of shoulder is important. I prefer the shoulder that is filled with flesh!"


[redacted]

انركب لولب مثلكم والا تركبن اصطناعي مثلنا هيك ضل بس

Translation: "It remains only for you to install an artificial male organ or for men to install an IUD."

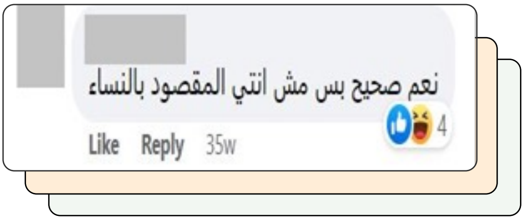


[redacted] [redacted]

انداري مستحي احكي شو ضل غير آنكم تركبوا . ....

Translation: "I'm ashamed of what to say; I think the only thing left for women now is to "hump" men."

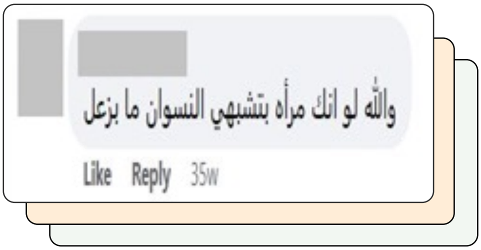
B. Misogyny: Some men users hinted at Bani Mustafa having more masculine characteristics than the feminine characteristics that women usually possess; examples:



[redacted] [redacted]

نعم صحيح بس مش انتي المقصود بالنساء

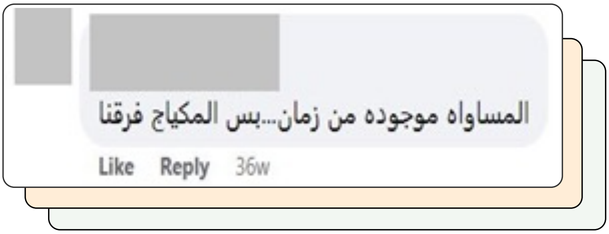
Translation: "Yes, that's correct, but you are not included in this because this is only directed towards women."



[redacted] [redacted]

والله لو انك مراه بتشبهي النسوان ما بزعل

Translation: "If you were a woman that actually looked like women, I would not be as angry."



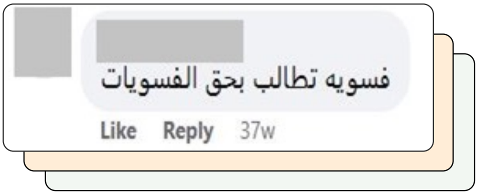
[redacted] [redacted]

المساواه موجوده من زمان...بس المكياج فرقنا

Translation: "Gender equality exists from a very long time ago, but make-up has created a distinction between genders!", inferring that makeup is the only feminine factor associated with Bani Mustafa.

4. Personal Abuse:

Mrs. Bani Mustafa was abused with offensive descriptions targeting her position in public life and in her personal life; examples:

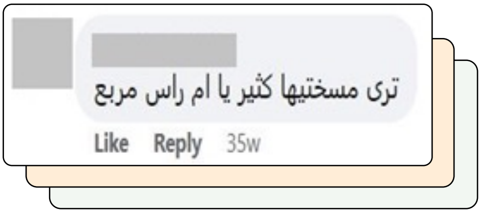


[redacted]

فسويه تطالب بحق الفسويات

Like Reply 37w

Translation: An insult directed at Bani Mustafa and women demanding gender equality by using a term that means farting as a replacement for the word "feminist".

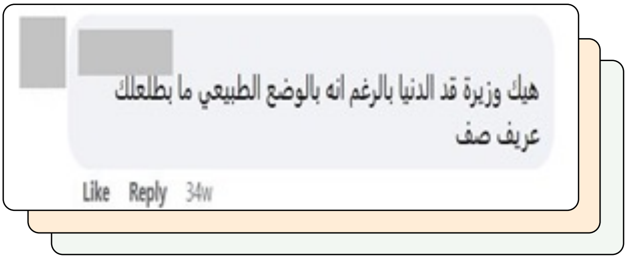


[redacted]

ترى مسختيها كثير يا ام راس مربع

Like Reply 35w

Translation: "You have crossed the limits in disrespecting the customs and traditions of society, you square-headed person."



[redacted]

هيك وزيرة قد الدنيا بالرغم انه بالوضع الطبيعي ما بطلعلك عريف صف

Like Reply 34w

Translation: "You got the position of minister, although, in normal situations, you would not be able to get the position of class corporal in a school."

Research Limitations

The report has some limitations that need to be acknowledged. One of the major limitations is that some comments were deleted or hidden, which prevented the SMM team from analysing the total interactions. Additionally, the harmful and offensive language used in the comments had a negative impact on the research team and social media users in general. After completing the analysis of each report, we acknowledged the SMM team's hard work by granting them a well-deserved day off. This break served as an opportunity for them to recover from the challenging content they encountered during their work. Furthermore, we consistently encouraged open communication among team members, urging them to discuss their experiences so as to help release any lingering thoughts or emotions.

It is important to note that the study focused only on a specific case study, and may not be representative of the overall online discourse on gender issues in Jordan. Further research is needed to provide a more comprehensive understanding of the topic.

Another limitation of the study was the manual data-collection process, which was time-consuming and limited the amount of data that could be analysed. However, the use of the Export Comments tool helped mitigate this to some extent. By utilising the Export Comments tool, the team was able to retrieve a larger dataset of over 3,500 comments, including nested comments.

Although the data collection process was facilitated by the Export Comments tool, the subsequent step of labeling still relied on manual efforts. The team manually reviewed each comment and applied labels based on the presence of hate speech or non-hateful content. This manual labeling process ensured accuracy in categorising the comments and identifying hate speech.

While the Export Comments tool enabled the collection of a larger dataset, it's important to acknowledge that the manual labeling process might have introduced limitations. Due to the manual nature of the labeling, there is a possibility that some relevant comments were missed or mislabeled, which could have influenced the overall findings of the study.

Recommendations:

Civil Society:

- Provide support and resources for SMM teams to manage and cope with the emotional impact of monitoring hate speech and disinformation, including training on self-care and mental health.
- Conduct further research on the impact of hate speech and disinformation on gender-based issues, particularly in countries with high levels of gender inequality, to inform the development of effective policies and programmes.
- Prioritise research on gendered disinformation in the MENA Region, by conducting further research on the impact of gendered disinformation.
- Increase public awareness and education about the harmful impact of hate speech and disinformation on society, and particularly on women and girls.
- Develop and implement gender-sensitive policies and programmes that address gender inequality and discrimination in various areas, such as education, employment, and political participation.
- Promote dialogue and open communication between different groups and individuals with diverse perspectives on gender issues, to foster greater understanding and respect for different viewpoints.

Media:

- Encourage media outlets to provide more balanced and diverse coverage of gender issues, including featuring the voices and experiences of women from different backgrounds and perspectives.

Religious Institutions:

- Engage with religious leaders and scholars to promote a more nuanced and progressive understanding of gender issues within the context of religious teachings and traditions.

Spotlights:

1. The Words Matter network aims to create and measure the impact of different interventions to counter information manipulation and online violence on social media platforms. Researchers believe that pre-bunking has proven successful in raising awareness about main tactics and methods of manipulating social media. The Maharat Foundation has been working to raise the resilience of voters to online mis/disinformation, by building an inclusive information literacy campaign driven by the lessons learnt from their monitoring of the parliamentary elections in 2022, and by preparing the Lebanese citizens to deal with various online threats that can be expected during the campaign for the local elections in 2023.

2. Institut de Presse et des Sciences de l'Information (IPSI) has conducted research on the generation of political disinformation in Tunisia. [21] In the second spotlight, the authors present their work and their conclusions on improving the Tunisian online environment and making it less susceptible to disinformation.

4. Maharat “SPOT IT, CHECK IT, STOP IT”

An inclusive information literacy awareness campaign

Media and information literacy is considered an essential tool for spreading awareness and necessary information to society, and especially to journalists and journalism students, to help them identify, detect and learn how to fact-check mis- and disinformation, and to distinguish its different types. False and misleading information accounts for a large part of all information in circulation, according to several reports produced by the Maharat Foundation in recent years. One such report, “Political Propaganda and Information Manipulation on Social Media during the Lebanese Parliamentary Elections 2022”, stated that 49.5 per cent of the political discourse in online circulation was based on manipulation of feelings and emotional political propaganda, instead of being based on factual information or legal arguments to back claims or counter those of opponent parties. Likewise, hate speech and violence against women in political and public life formed a large part of the public discourse.

An unbalanced and distorted informational environment, where engaging in political and public discussions becomes unsafe, particularly amid elections, with a lack of

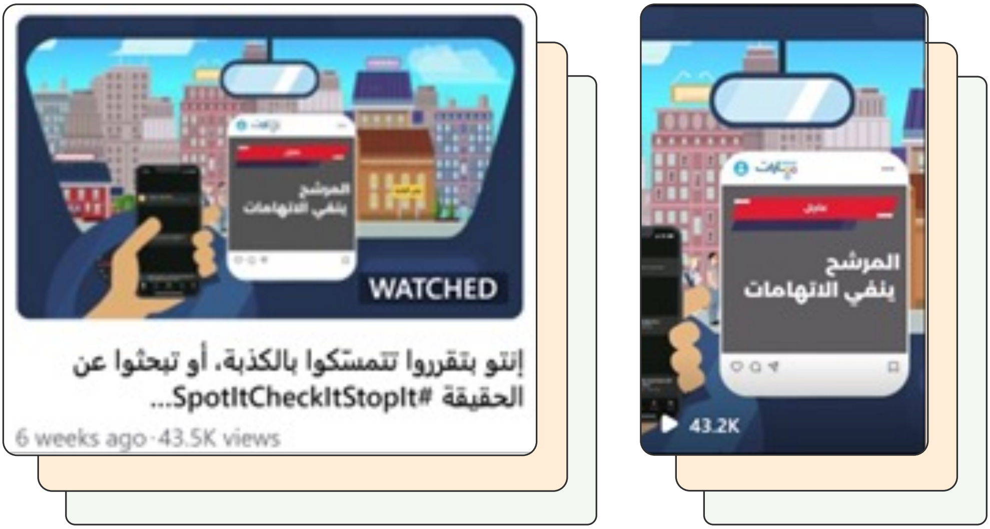
[21] http://www.mourakiboun.org/storage/bKHaeOXHhFABcE04yERwVq7WOdmPkViQlpcKijBL.pdf

independent media, large-scale fact-checking, and digital information, and media literacy initiatives, contributes to the uncontrolled spread of disinformation.

Therefore, based on our findings and the evidence-based data we collected through our monitoring during the period before, during and after elections, the Maharat Foundation worked on providing resource materials to unpack various aspects of information circulation and political and public discourses, to help journalists and journalism students gain a better understanding of the informational environment, especially in the run-up to this year's municipal elections, through the ongoing online information literacy campaign "SPOT IT, CHECK IT, STOP IT", launched in March 2023.

Campaign launch:

The campaign's launch was based on an animated short video of 1m:09s on Instagram and Facebook.



إنتو بتقرروا تتمسكوا بالكذبة، أو تبحثوا عن
الحقيقة #SpotItCheckItStopIt...

6 weeks ago · 43.5K views

Figure 01: Screenshots from Facebook and Instagram of the launch video for the campaign

The video brought together the four topics we are working on – false information, political propaganda, hate speech, and violence against women in politics. The video shows the influence of publishing false information, intentionally or unintentionally, with the aim of harassing a candidate or attacking them, in order to put them out of the electoral race, or any information that may lead to violence that may affect a man or woman working in the public field (politician, journalist, activist, civil servant, etc.).

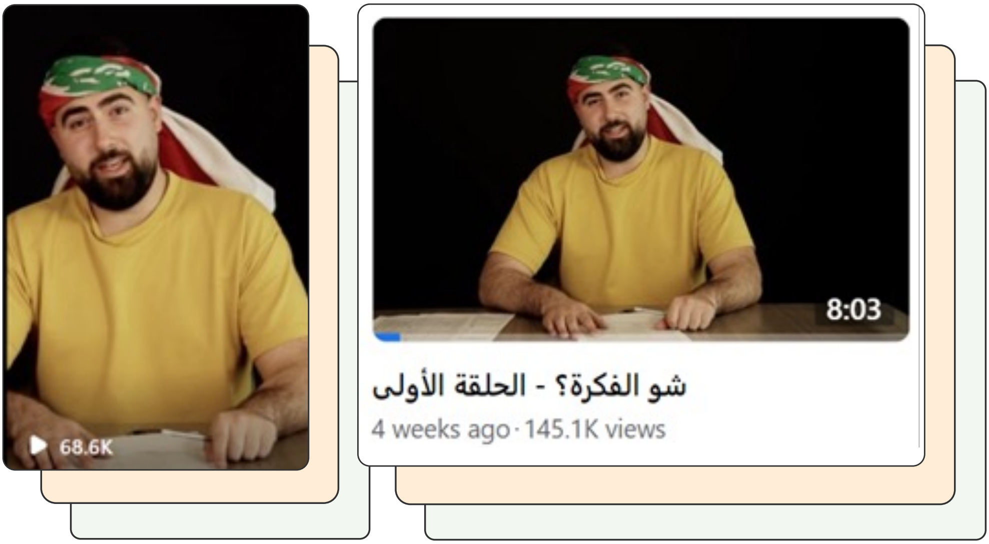
The goal of the video is to urge the targeted audience (journalists and journalism students) to deal with the news more consciously and accurately, and to always verify the validity of any circulating news before publishing it, in order to avoid misinformation and misleading the public opinion or harming anyone.

Political propaganda:

As part of creating new content for the Maharat Foundation, a series titled "Chou elfekra?" or "What's the Idea?" was produced.

The first episode of this series was about political propaganda. The Hosts Bilal Yassine and Laura Rahal discussed the concept of political propaganda, its methods, and three strategies that can be adopted to avoid it. It also featured a satirical skit symbolising the political propaganda used by politicians in their speeches during their electoral campaigns.

This first episode generated high engagement, reach and views on both Instagram and Facebook.



شو الفكرة؟ - الحلقة الأولى

4 weeks ago · 145.1K views

Figure 02: Screenshots from Facebook and Instagram of the first episode of the "Chou lfekra" series, with Bilal Yassine, one of the hosts.

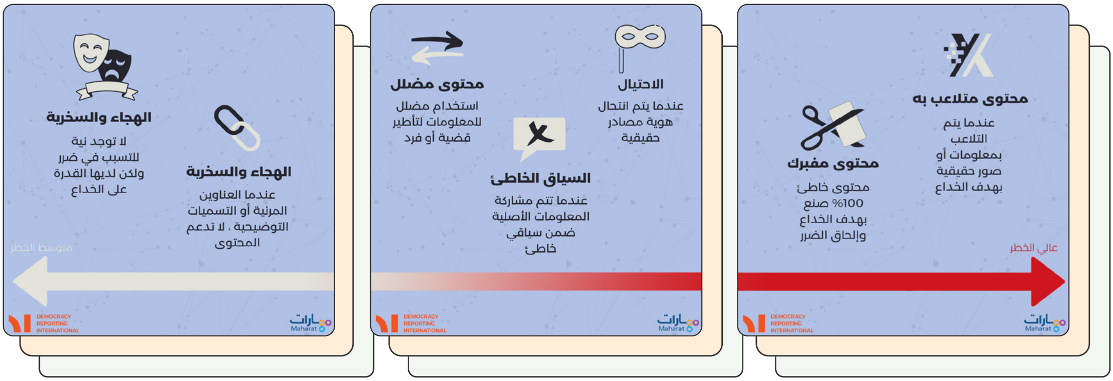
A carousel post was also shared on social media explaining the differences between the methods used in political propaganda: Emotional grooming vs. allegations. (Instagram, Facebook)

Fact-checking:

Previous studies and reports produced by the Maharat Foundation during the 2022 parliamentary elections identified patterns of false news topics that can circulate during and after elections.

Based on this, and anticipating the upcoming expected municipal elections, the Foundation produced two carousel posts. The first post listed thematic areas that could be part of disinformation campaigns before the municipal elections (Instagram, Facebook), while the second listed thematic areas that could be part of disinformation campaigns after the municipal elections (Instagram, Facebook).

The Maharat Foundation also posted a carousel graphic explaining the different tactics of disinformation and the levels of harm they do to society (Instagram, Facebook).



الهجاء والسخرية
لا توجد نية
للتسبب في ضرر
ولكن لديها القدرة
على الخداع

الهجاء والسخرية
عندما العناوين
المرئية أو التسميات
التوضيحية ، لا تدعم
المحتوى

متوسط الخطر

الاحتيال
عندما يتم انتحال
هوية مصادر
حقيقية

محتوى مضلل
استخدام مضلل
للمعلومات لتأطير
قضية أو فرد

السياق الخاطئ
عندما تتم مشاركة
المعلومات الأصلية
ضمن سياقي
خاطئ

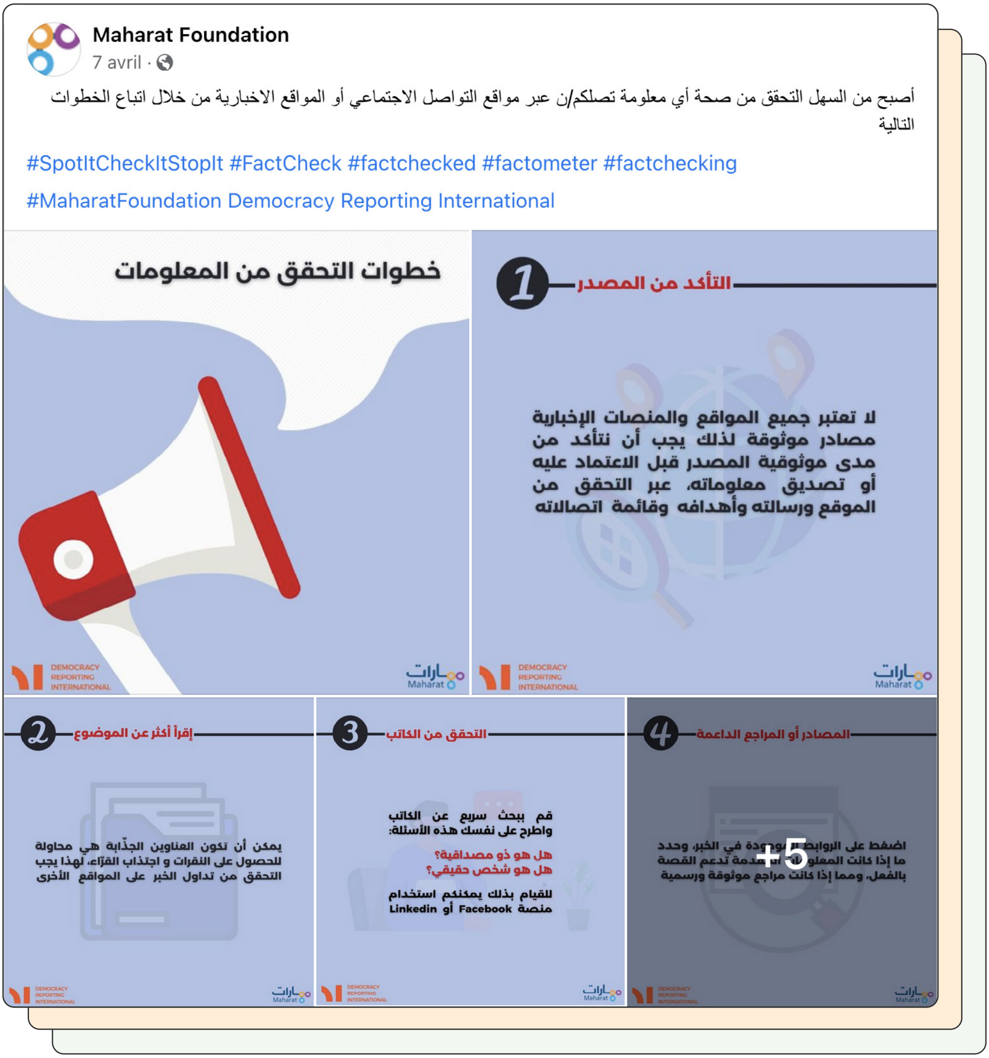
محتوى متلاعب به
عندما يتم
التلاعب
بمعلومات أو
صور حقيقية
بهدف الخداع

محتوى مفبرك
محتوى خاطئ
100% صنع
بهدف الخداع
وإلحاق الضرر

عالي الخطر

Figure 03: Types of false information and the level of harm they do to society.

We also illuminated the seven steps of fact-checking, as adopted by Maharat News Fact-o-meter (Instagram, Facebook).



Maharat Foundation
7 avril ·

أصبح من السهل التحقق من صحة أي معلومة تصلكم/ن عبر مواقع التواصل الاجتماعي أو المواقع الاخبارية من خلال اتباع الخطوات التالية

#SpotItCheckItStopIt #FactCheck #factchecked #factometer #factchecking #MaharatFoundation Democracy Reporting International

خطوات التحقق من المعلومات

1 التأكد من المصدر

لا تعتبر جميع المواقع والمنصات الإخبارية مصادر موثوقة لذلك يجب أن نتأكد من مدى موثوقية المصدر قبل الاعتماد عليه أو تصديق معلوماته، عبر التحقق من الموقع ورسالته وأهدافه وقائمة اتصالاته

2 إقرأ أكثر عن الموضوع

يمكن أن تكون العناوين الجذّابة هي محاولة للحصول على النقرات و اجتذاب القرّاء، لهذا يجب التحقق من تداول الخبر على المواقع الأخرى

3 التحقق من الكاتب

قم ببحث سريع عن الكاتب واطرح على نفسك هذه الأسئلة:

هل هو ذو مصداقية؟
هل هو شخص حقيقي؟

للقيام بذلك يمكنكم استخدام منصة Facebook أو Linkedin

4 المصادر أو المراجع الداعمة

اضغط على الروابط الموجودة في الخبر، وحدد ما إذا كانت المعلومات المقدمة تدعم القصة بالفعل، ومما إذا كانت مراجع موثوقة ورسمية

Figure 04: A screenshot from Facebook of the seven steps of fact-checking adopted by Maharat News Fact-O-Meter.

Awareness around the municipal elections:

As part of its awareness-raising campaign, the Maharat Foundation produced an entertaining, yet informative quiz on Instagram and Facebook stories, to let our audience share their thoughts about the importance of the municipal elections and the participation of women, as well as letting them test their knowledge about the laws and regulations governing the municipal elections.

Violence against women in public and political life:

The second episode of "Chou elfekra" tackled the topic of violence against women in political and public life.

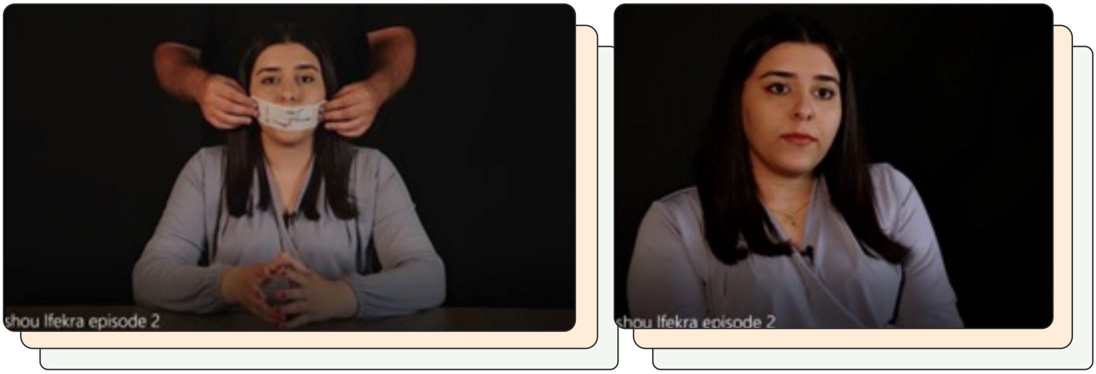


Figure 05: A screenshot from the second episode of "Chou lfekra", showing one of the programme's hosts, Laura Rahal

The Hosts Laura Rahal and Bilal Yassine discussed the phenomenon of violence against women in political and public life, and also explained the different types of violence against women heir motivations. The episode encourages women to raise their voices, by reporting any kind of violence they have been exposed to in public and political life through the hotline set up by the Maharat Foundation.

Campaign Impact:

Through this media information literacy campaign, the Maharat Foundation aimed to make an impact on its audience on two levels. On the first level, this was to reach our main audience of around 700 journalists and student journalists through targeted newsletters and Maharat's website, to raise their awareness of the importance of fact-checking and give them a better understanding of the political propaganda and hate speech that target different communities and categories of society, in order to create a more balanced informational environment. On the second level, this targeted the Maharat Foundation's general audience, to help them develop a critical sense of information consumption through Maharat's social media presence. For example, the video disseminated during the campaign about political propaganda reached 68,600 viewers on Instagram and 145,100 on Facebook.

Despite the postponement of the municipal elections to an unspecified date, this campaign will continue, and new content will be published soon, on the Maharat Foundation website, where all materials related to the campaign are available.

Manufacturing Political Disinformation in Tunisia

Interview with Institut de Presse et des Sciences de l'Information- IPSI

Conducted by **Hervé de Baillenx**

The Words Matter team conducted an interview with Tunisia's Institute of Journalism – Institut de Presse et des Sciences de l'Information – IPSI, [22] to present their recently published research paper.

The full research document can be found here. [23]

Hello, you are the authors of a research paper recently published by IPSI on the manufacturing of political disinformation in Tunisia. Could you introduce yourselves?

I am Dr. Maroua Ben Becha, assistant professor at the Institut de Presse et des Sciences de l'Information;

I am Dr. Khalil Jelassi, Dr. of information and communication sciences, and digital media expert.

[23] http://www.mourakiboun.org/storage/bKHaeOXHhFABcE04yERwVq7WOdmPkViQlpcKijBL.pdf

Could you tell us what the study is about?

Tunisia is a country that has been undergoing a political transition since 2011. It has had several significant political moments, the most recent of which was 25 July 2021 [the constitutional referendum], which radically changed political life. [24] There is a great deal at stake in manipulating public opinion, especially now that Tunisia is experiencing political polarisation.

This is why we produced this study, entitled "Manufacturing Political Disinformation in Tunisia". It has two objectives. One is to conceptualise disinformation in Tunisia in relation to other regional or international contexts. The other is to identify the strategies used to give credibility to this disinformation. The study has a theoretical component and an empirical component.

What methodology did you use?

The study mixed several techniques. We chose Facebook because Tunisians use this medium as a source of information and news. We analysed disinformation content on Facebook pages. We then combined the content we found with cases selected from Tunifact's [25] content verification pages and other fact-checkers, and we also used studies carried out in the Tunisian context by authors such as LabTrack, Inkyfada, and DRI-Atide. This was a meta-analysis, a study of studies. Then we identified our own corpus to draw more conclusions.

Whether it's the general public, the government, the media, or political parties, everyone is responsible for disinformation. The aim of our study was not to find out who the players are, but to decipher the tactics and the strategies of manipulation.

You make use of a key concept, “informational disorder”. What is this?

Information disorder is the opposite of an ordered system of information. It is a state of society in which sources of information are scrambled by different players, in which citizens have no access to verified information.

Can you give a concrete example of such informational disorder?

In our research, we chose a news item. It was the Tunisian-Moroccan diplomatic incident. [26] We chose a collection of several Facebook pages that dealt with this subject from different political parties – pro-Saïd, pro-Abir Moussi, [27] and pro-Ennahda. There were many fakes, for example, repeated reports about Attijari Bank's headquarters [28] leaving Tunisia.

Is this informational disorder organic, or do you see a coordinated campaign to influence public opinion and users?

Some behaviour is natural. There are people who spread this false information because they believe in a crisis, but, above all, the large part involves strategies that try to manipulate public opinion.

[24] On this date, the Tunisian president suspended (and later dismissed) the Parliament, concentrating in his hands both executive and legislative power.

[25] Tunifact is the factchecking platform of the Tunisian Union of Journalists.

[26] The two countries recalled their ambassadors after Tunisian President Kais Saied met with the leader of the Polisario Front at the eighth Tokyo International Conference on African Development (TICAD 8), held in Tunis on August 27 and 28, 2022.

[27] She is one of the main opposition leaders.

[28] A Moroccan bank

What are the main findings?

The main finding is that **disinformation has become systemic**. We are no longer talking about practices scattered here and there. It has become a weapon of political persuasion, and that is dangerous.

One of the main strategies specific to the Tunisian context is that **these pages imitate a journalistic format** and journalistic style in the promotion of content, for example, by using journalistic jargon such as "according to our sources". This is designed to lend credibility to these productions. **Humour** is also one of the greatest vectors of disinformation, as satire and parody are used to manipulate. These two strategies are inseparable. You make sure that the content, whether a video, an image or cartoon, or a simple post, acquires a certain form of credibility, so that it will be believed and shared.

There is also the **use of Facebook Live**. This is very important, because Facebook live posts have a considerable organic reach compared to the publication of a standard post.

We identified catalysts. They are (i) filter bubbles; (ii) algorithms that tend to favour certain disinformation content; (iii) political polarisation; (iv) and, finally, content that is emotionally charged. They contribute to intensifying the virality of disinformation.

This study may have been biased by the 25 July 2021 event, after which everything became polarised. Either you are for this political project, or you are against it. All phenomena – social, economic, even cultural events like the Francophonie summit, have been subject to political disinformation, and all this is the result of the division in society.

Facebook has become a battleground for political parties, where any crisis is exploited. It is a battleground for political parties, an arena for the exercise of power and influence, where any crisis is exploited politically, economically, socially, and culturally.

Do you have any recommendations?

We have three main areas of recommendations.

1- We have had 11 governments since 2011, resulting in no clear public policy for the media sector. This study calls for **public policies aimed at protecting the media and strengthening their role in society**, and, above all, by creating powerful and influential public media hubs to act as a counterweight to this ocean of disinformation. A credible media is one that is economically powerful, independent, and serves the general interest. Unfortunately, in Tunisia we tend to go back to the old model, which relied on government media.

2- The second recommendation is **education for this digital age**, starting in primary school as a subject in its own right. It is essential to teach modules on distortion, on how to make good use of media and mobile terminals.

3- The final recommendation is to open **a national debate**, setting up a council bringing together experts on information science, sociology, and education, and government representatives.

Any final message?

Online disinformation is a scourge in many countries, but the challenge in Tunisia is twofold: (i) There needs to be a collective awareness; and (ii) Unfortunately, this kind of initiative always comes from civil society, and not at all from the state or the media.

We would like to supplement this work with other studies. We covered production strategies, and now we want to complete the picture with how users consume information.

About Words Matter

Contact: wordsmatter@democracy-reporting.org

DRI has been increasingly active in the field of social media monitoring (SMM) since 2017, strengthening local capacities to monitor social media during elections, sharing information and evidence gathered in different countries, bringing together expert organisations, producing methodologies, and informing public and expert debate.

Within the framework of the "Words Matter" project, DRI and its partners seek to contribute to strengthening the safeguarding of democratic processes and societies' resilience to online disinformation and hate speech in the MENA region.

DRI works with partner organisations from four countries (Jordan, Lebanon, Sudan, and Tunisia), strengthening local capacities to monitor and analyse online disinformation and hate speech during key national democratic processes, while building a regional network to allow for comparative analysis and peer learning.

"Words Matter" aims to achieve the following objectives:

- Capacity-building for project partners **to acquire institutional skills to design sound social media monitoring methodologies**, to effectively monitor disinformation and hate speech online, and to enhance evidence of the impacts of disinformation and hate speech online on civic or political participation and human rights.
- **Enhanced multi-stakeholder and regional engagement** to advocate against and combat online disinformation and hate speech, through a civil society network, as well as through continuous exchanges on transparent regulations.; and
- In the countries of project partners, **improved awareness and resilience of civic target groups**, and concrete action by decision-makers to transparently combat online hate speech and disinformation.

About the Digital Democracy Program

Contact: info@democracy-reporting.org

DRI's Digital Democracy (DD) programme protects online democratic discourse by exposing disinformation, manipulation and hate speech, strengthening the capacity of CSOs for monitoring and advocacy, and ensuring appropriate and evidence-based responses from governments and tech companies.

DRI is well-positioned to address online threats and disinformation, due to its research on manipulated media content, deepfakes as potential disinformation tools, and its current focus on identifying new potential threats and emerging technologies in this field. As part of our diverse toolbox, we have, for example, integrated machine learning models to help us identify emerging trends in the disinformation space. Our work on information manipulation is also complemented by analysing and publishing guides on gender-based under-representation and harassment online.

An important activity within the DD programme for exposing and fighting hate speech and disinformation is social media monitoring (SMM). SMM is the objective analysis of democratic discourse and political actors on social media platforms. This is far more complex than traditional media monitoring, with a myriad of actors and content, combining official democratic institutions (e.g., political parties, politicians, media) and unofficial actors (e.g., individuals, political influencers, partisan groups). This is why DRI published the Digital Democracy Monitor Toolkit, the first social media monitoring methodology that helps civil society, journalists, and academia to research social media and democracy.

Our methodology was tested and used for conducting social media monitoring in 12 countries (including Germany, Libya, Myanmar, Nigeria and Sri Lanka), focusing on disinformation, hate speech and political advertising before, during and after the elections. By using a holistic approach to analyse social media, our toolkit engages with disinformation and hate speech by looking at the message or content, the active messengers, and the messaging, thus both the forms and the channels of distribution.

Based on the findings of our SMM, we have advocated for the implementation of the European Democracy Action Plan (EDAP) commitments, which could strengthen the fight against disinformation at the EU level and contribute to the debate about content-ranking systems, a major challenge when it comes to the dissemination of dis/misinformation. DRI has also lobbied for the implementation of the EU's Digital Service Act, a potential milestone in the effort to increase accountability across social media platforms. In launching the Arabic version of the SMM toolkit, we hope to empower the MENA region in the same way.

About DRI

Contact: info@democracy-reporting.org

Democracy Reporting International (DRI) is an independent organisation dedicated to promoting democracy worldwide. We believe that people are active participants in public life, not subjects of their governments. We strengthen democracy by supporting the institutions and processes that make it sustainable, and work with all stakeholders towards ensuring that citizens play a role in shaping their country. Our vision is grounded in globally agreed upon principles of democracy, stemming from the democratic governance championed by the United Nations and international law.

DRI's work focuses on five key themes of democracy: Justice, Elections, Local Governance, Digital Democracy and Human Rights. By working at both the national and local level, we use five intervention approaches in our projects: awareness-raising, capacity-building, fostering engagement between different stakeholders, supporting the building of democratic institutions, and advising on the drafting and implementing of policies and laws.

DRI's work is led by a Berlin-based executive team and supervised by an independent board of proven democracy champions. DRI maintains country offices in Lebanon, Libya, Tunisia, Pakistan, Myanmar, Sri Lanka and Ukraine. Through our networks of country offices and partners, we are in a unique position to track, document, and report developments and help make tangible improvements on the ground.

About DRI Partners

Al-Hayat Center for Civil Society Development: is a non-governmental civil society organization founded in 2006. The center has expanded to become one of the leading NGOs in Jordan. Al-Hayat's overall mission is to promote accountability, governance, public participation, and tolerance in Jordan and the region within the framework of democracy, human rights, the rule of law, and gender mainstreaming in public policy and actions.

Jordan Open Source Association (JOSA): is a non-profit organization based in Amman, Jordan. The association is among the few non-profits registered under the Jordan Ministry of Digital Economy and Entrepreneurship. JOSA's mission is to promote openness in technology and to defend the rights of technology users in Jordan. JOSA believes that information that is non-personal – whether it's software code, hardware design blueprints, data, network protocols and architecture, content – should be free for everyone to view, use, share, and modify. JOSA's belief also holds that information that is personal should be protected within legal and technological frameworks. Access to the modern Web should likewise remain open.

Lab'TRACK: is a laboratory for monitoring, analysis and reflection on political disinformation phenomenon on social networks, in particular the Facebook network. The laboratory is a collaboration between Mourakiboun and IPSI.

MOURAKIBOUN: Mourakiboun is a domestic electoral observation network that was launched in 2011 and is today a key player in this field with multiple national and international partners. Since 2014, Mourakiboun has been diversifying its actions by adding accountability of public services and support to the Tunisian decentralization process to its portfolio. Mourakiboun has a network of over 100 volunteers in all regions of Tunisia and excellent access to local structures and stakeholders. Mourakiboun has adopted an IT approach to its activities, thereby increasingly reaching Tunisian youth. During the 2014 and 2019 presidential elections, Mourakiboun conducted social media monitoring activities focused on the interactions of FB users with the speeches of candidates during electoral campaigns.

Institut de Presse et des Sciences de l'Information (IPSI): was established in 1967 and became a non-departmental public institution enjoying financial autonomy and legal personality in 1973. The Institute is known as Tunisia’s leading university for the education of journalists and media workers. IPSI’s research in the field of information and communication sciences has been met with international acclaim. IPSI has a network of national (INLUCC, HAICA, UFP) and international partners (Deutsche Welle Akademie, UNESCO, UNDP, Article 19 among others). Through this cooperation, IPSI provides specialized training sessions and hosts experts and internationally renowned speakers to introduce students to innovative practices in the field of communication.

MAHARAT: a women-led, Beirut-based organization, working as a catalyst, defending and advancing the development of democratic societies governed by the values of freedom of expression and respect for human rights.

Maharat advances the societal and political conditions that enhance freedom of expression and access to information, both online and offline. Maharat engages and equips a progressive community in Lebanon and the MENA region with the skills and knowledge necessary to create change.

Sudanese Development Initiative (SUDIA): In light of the ongoing armed conflict in Sudan, our valued partner from Sudan, The Sudanese Development Initiative (SUDIA), has regrettably been unable to continue their participation in the project. We deeply appreciate the significant contributions they made during their active involvement. While they are no longer with us due to these challenging circumstances, their dedication and expertise have left a lasting impact on the project's progress. As we move forward, we honor their commitment and extend our hopes for a peaceful resolution to the conflict in Sudan.

Glossary:

TERMS	DEFINITIONS
Code-driven annotation	A data-annotation process that uses Python code to automatically label or tag data, based on predefined rules, algorithms, or patterns.
Cohesive message	A well-structured and unified communication that presents a clear and consistent theme or idea. It is a message that flows logically and coherently, with each element reinforcing and supporting the central concept.
Conspiracy theories	Explanations or beliefs that suggest a hidden and secretive group or organisation is responsible for orchestrating or controlling significant events, actions, or outcomes. These theories often propose that powerful entities, such as governments, corporations, or influential individuals, are engaged in secretive plots to deceive the public and manipulate events for their own benefit.
Coordinated campaigns	Organised efforts in which individuals or groups work together in a planned and synchronized manner to achieve a common goal. These campaigns involve cooperation, communication, and strategic coordination among the participants, to maximise their impact and effectiveness.
Data annotation	The process of adding metadata or tags to raw data to make it understandable and usable for data analysis.
Data augmentation	Techniques used to generate new synthetic datasets, based on original data.
Death threats	Communications that include clear death threats.
Deceptive Links	Hyperlinks or URLs that are intentionally designed to mislead or trick users into clicking on them. These links are created to appear as something other than what they actually lead to, often with the goal of directing users to malicious websites or deceptive content.

TERMS	DEFINITIONS
Defamation	Communications that aim to defame someone's reputation.
Disinformation	False information that is spread with the intent to mislead and do harm, and in an organised way.
Domestic	Messages originating from within a country's own borders.
Election day	Election day, also known as polling day, is the day when citizens of a country or a region vote to elect their representatives.
Election silence/ Pre-election silence	Election silence, also known as pre-election silence or electoral silence, is the period of time before an election during which political campaigning is prohibited. This period of time is established by law and is intended to give voters a chance to reflect on the issues and make a decision without the influence of last-minute campaign materials or statements.
F1	A metric used to evaluate a machine-learning model by measuring its accuracy.
False information	Completely untrue information, with a lack of any factual basis.
False negative (FN)	An outcome where a model incorrectly predicts the negative class (i.e., when the model fails to recognise that content contains hate speech and labels it as hate speech-free)
False positive (FP)	An outcome where the model incorrectly predicts the positive class (i.e., when the model labels hate speech-free content as hate speech).
Hand-picked selection	A manual choice of a group of accounts or sources, intentionally selected by individuals, with careful consideration and for a specific purpose, rather than being chosen randomly or automatically.
Hate speech	A form of communication, whether written, spoken, or expressed through other means, that promotes violence, discrimination, hostility, or prejudice against individuals or groups based on attributes such as race, ethnicity, religion, gender, sexual orientation, nationality, or disability.

TERMS	DEFINITIONS
Investigation	Systematic and thorough examination or study conducted to gather information, analyse evidence, and uncover facts or truths about a particular subject.
Jupyter notebooks	Interactive computing environments that enable users to execute Python code, perform data analysis, and create documentation within a single interface. They provide a convenient way to combine code snippets, visualisations, and explanatory text.
Left-wing parties	Political parties that advocate for progressive and liberal policies, often focusing on social equality, workers' rights, and government intervention in the economy.
Machine learning	A branch of artificial intelligence (AI) and computer science that focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.
Mal-information	Any kind of communication that involves leaks, doxing, or the release of private information. This can be true or false.
Misinformation	False information that is spread, regardless of the intent to mislead.
Misleading information	Content, data, or statements that are intentionally or unintentionally inaccurate, deceptive, or misleading. This information can be presented in various forms, such as written text, images, audio, or video, and is designed to deceive, manipulate, or mislead the audience. Misleading information may contain partial truths, exaggerations, or omissions that create a distorted or biased view of a particular subject.
Misogyny	Communications that focus on gender stereotypes and gender-based prejudice.
Narratives	Structured or coherent accounts of events, experiences, or information presented in a story-like format. A way of organising and conveying information to create meaning and understanding for the audience.
Networks	Groups of Facebook pages that have been discovered to be strategically coordinating their content and activities.

TERMS	DEFINITIONS
Online gender-based violence (OGBV)	Action online by one or more people that harms others based on their sexual or gender identity or by enforcing harmful gender norms.
Page morphing	A deceptive practice that can be used to circumvent Facebook's policies and mislead users about the true nature or purpose of a page.
Pattern	A repeated and recognizable arrangement or sequence of elements, shapes, symbols, or behaviours. It refers to the regular and consistent recurrence of certain features or characteristics that can be identified and distinguished from other elements.
Political affiliation	An individual's formal or informal association with a particular political party, ideology, or group. A person's alignment or identification with a specific set of political beliefs, values, and policy positions.
Political content	Content that refers to a candidate, political party, elected or appointed government official, election, referendum, ballot measure, legislation, regulation, directive, or judicial outcome.
Political polarisation	The widening ideological and emotional divide between different political groups or individuals with contrasting viewpoints. It occurs when people's political beliefs, values, and attitudes become more extreme and sharply divergent, leading to a reduced willingness to find common ground or compromise on issues.
Power BI	A business analytics service and data visualisation tool developed by Microsoft. It allows users to connect to various data sources, prepare and transform the data, and create interactive and visually appealing reports and dashboards.

TERMS	DEFINITIONS
Pre-elections	Pre-elections refers to the period before an election is held, during which political campaigns, discussions, and preparations take place. The pre-election period sets the stage for the actual election and can have a significant impact on the outcome of the election.
Programming logic	The set of rules, principles, and techniques used in coding to solve problems and achieve specific outcomes. It involves creating algorithms and logical sequences of instructions that guide the computer's operations, ensuring that tasks are performed accurately and efficiently.
Sarcasm	Communication that uses irony, mockery, or satirical remarks to convey a different meaning than literal interpretation.
Sexual harassment	Communications that contain expressions of sexist and sexual harassment.
Threats of physical violence	Communications that threaten the infliction of physical violence.
Twitter API	A set of programmatic endpoints that can be used to understand or build the conversation on Twitter. This API allows one to find and retrieve, engage with, or create a variety of different resources, including the following: Tweets. Users. Spaces.
Weaponisation of mockery	The intentional and strategic use of ridicule, sarcasm, or derision as a tool to undermine, discredit, or harm individuals, groups, or ideas. It involves using humour or satire not merely for entertainment or light-hearted banter, but as a means to attack or belittle a target, often with the aim of achieving specific political, social, or ideological objectives.

DEMOCRACY
REPORTING
INTERNATIONAL

Democracy Reporting International (DRI) was founded in 2006 by an international group of experts on democratic governance and elections.

DRI works on research and analysis to direct engagement with partners on the ground to improve democratic structures and safeguards across the countries where we work.

Elbestraße 28/29 12045 Berlin, Germany
info@democracy-reporting.org
wordsmatter@democracy-reporting.org
www.democracy-reporting.org/