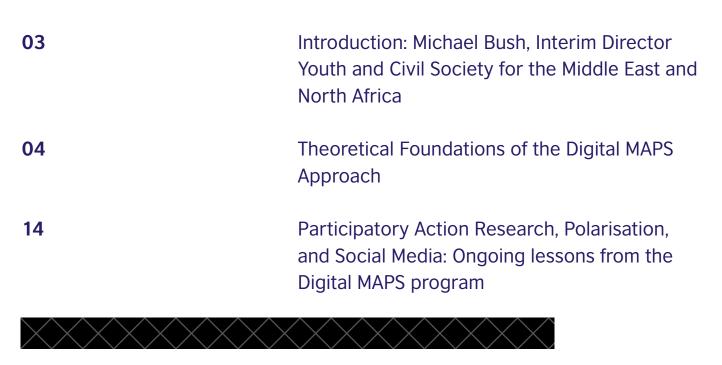


**DMAPS Reports** 







121	Research Reports:
121	Iraq
129	Lebanon
$\times \times \times \times \times \times \times$	$\longleftrightarrow$

The British Council's vision is a more peaceful and prosperous world built on trust. We support peace and prosperity by building connections, understanding and trust between people in the UK and countries worldwide through arts and culture, education and the English language.

Good Cultural Relations can't ignore the ever-growing role of information and communication technologies, especially social media. These developments have brought with them the potential for huge benefits, including the proliferation of Digital Cultural Relations, defined here as "cross-cultural practices through digital and networked technologies, including the internet, mobile devices, and social media channels"<sup>1</sup>. However, whilst many such practices are designed to support mutual understanding and peaceful co-existence, others (either purposefully or unwittingly), do the opposite.

What you'll find in this digital brochure is an explorative attempt to think through a Cultural Relations response to such negative digital and networked cross-cultural practices – i.e., a response to those online norms and practices that are damaging, rather than strengthening, people-to-people relations. Developed and delivered in 2021/2022, the Digital Media Arts for an inclusive Public Sphere (Digital MAPS) programme brought together three universities, a data science and software company, an international Digital Peacebuilding NGO, as well as 18 country-based media-arts initiatives, to explore local responses to affective polarisation – defined as the increasing dislike, distrust, and animosity towards those from other cultural or identity-based groups.

As you will read, through Digital MAPS, we worked with young leaders from the creative and media-arts sector, across eight countries in MENA (Libya, Palestine, Jordan, Yemen, Iraq, the Occupied Palestinian Territories (OPT), Tunisia and Syria), providing them with the skills, resources and technology to understand polarisation and depolarisation approaches, conduct their own social media analysis and on the basis of this design and deliver pilot interventions to undercut affective polarisation - whether it be centred on gender, ethno-sectarian conflict, intergenerational conflict or hate speech in general.

We hope the information contained here within, will be of interest to digital peacebuilding and digital cultural relations practitioners, policy makers and academics. More especially, we hope it can stimulate a conversation on the intersection between the two and the role of Cultural Relations in addressing the drivers that undermine it.

#### **Michael Bush**

Interim Director Youth and Civil Society for the Middle East and North Africa The British Council

1 More Europe. "Position paper on Digital Cultural Relations." More Europe-External Cultural Relations, June 2021, http://moreeurope.org/wp-content/uploads/2021/06/Policy\_Paper\_Digital\_Cultural\_Relations.pdf



Catherine Arthur and Stef Pukallus

Theoretical Foundations of the DMAPS Approach



#### Introducing Digital Media Arts for an inclusive Public Sphere (DMAPS)

Digital Media Arts for an inclusive Public Sphere (DMAPS) is a program that supports young leaders from creative and media arts organisations across eight countries (Libya, Palestine, Jordan, Yemen, Iraq, the Occupied Palestinian Territories (OPT), Tunisia and Syria) to conduct social media mapping in order to understand how issues of identity, social cohesion and inclusion play out on social media.

DMAPS is led by a consortium comprising Build Up, University of Sheffield, University of Manchester, the American University of Kurdistan and datavaluepeople in collaboration with the British Council. It uses participatory actionresearch (see position paper 2 for details) and is inherently participant-led – with 18 local organisations driving the research definition, analysis and intervention design, and delivery. Its aim is to promote openness and inclusivity in the digital public sphere. More specifically, the DMAPS program is an intervention that focuses on polarisation in online spaces and how to decrease such polarisation. It is of interest to both practitioners and academics engaged in peacebuilding, communication and the creative arts as well as local participatory-action research. The latter resonates with the British Council's understanding of cultural relations based on genuine reciprocity and mutual understanding, as well as on supporting local communities in their own capacity-building, building cultural relations within communities and beyond borders as well as approaches to local peacebuilding. Local peacebuilding emphasises that 'people at the local level (a) have the capacity to articulate, develop, and enact solutions to their own problems; (b) possess important knowledge and a better understanding about the complexities of the people, situation, context, and culture of the conflict; (c) maintain pre-existing, established relationships that enable them to network and engage with others in the peace process; (d) may be more likely to have the motivation to act, and act quickly, to resolve issues that directly impact them; and (e) have the staying power to sustain peace long after external actors have left' (Locally Driven Peacebuilding 2015: 2).

This unique collaboration between academics, practitioners, local organisations and researchers and the British Council is informed by theoretical research that was undertaken by consortium members prior to the program. In this first of three position papers we show how this theoretical research has fed into the DMAPS project, its approach to polarisation, communication and the arts as well as into the specific pilot interventions developed by the six clusters in the MENA region. The second paper outlines the participatory action research process and select findings from the social media mapping and analysis, and discusses opportunities, challenges, and lessons from this practitioner-led social media analysis process. The third position paper forthcoming focuses on the British Council's rationale for creating the framework for and funding the DMAPS collaboration. All three position papers recognise the fundamental importance of communication within societies and between societies.

# The importance of communication within societies and between societies

DMAPS is premised on the recognition that communication is fundamental to the fabric of civil society and its associative and cooperative capacity. It further acknowledges that social media, through its capacity to amplify, and to spread polarising content (including mis-/disinformation) at an unprecedented scale and speed, represents a risk to the type of inclusive public sphere required to build strong open and associative cultures – locally, nationally and internationally. This focus on building civil capacity and communicative resilience represents a significant shift away from mainstream interventions in the networked public sphere that narrowly focus on content moderation (including identifying and taking down hate speech) and digital literacy (understood as identifying and avoiding misinformation). Instead of considering social media as a technical entity operating in a vacuum DMAPS focuses on understanding and engaging with the complexity of how polarisation plays out on social media.

More specifically, DMAPS in terms of its communicative dimension<sup>1</sup> is underpinned by four areas of theoretical research:

# 1. Communication that divides

Communication (including the arts) can be divisive for civil societies. DMAPS is divided into two main phases: the first is very much focused on the ways in which communication on social media around a variety of issues can be divisive and polarising. Here DMAPS relies on research undertaken by Puig Larrauri and Morrison (2022) and Pukallus (2022), who have all examined what weaponised communication content can look like, albeit with different foci and to differing extents. Puig Larrauri and Morrison (2022) identify digital conflict drivers.<sup>2</sup> For them one of the affordances is the way in which digital technologies can create and spread divisive content – content which they categorise into hate speech, misinformation and disinformation. Hate speech<sup>3</sup> encompasses any communication that targets individuals or group for attack based on identity characteristics (e.g. race, gender, ethnicity, religion, etc.) and has become such an issue in the online space that it is now incorporated into various social media platforms' standards and rules. Misinformation and disinformation differ in that they are perhaps less explicit in their attack or ability to cause harm; misinformation is the dissemination of false information that could cause harm, especially in conflict contexts, but without the deliberate intention to do so by its agent. Disinformation, on the other hand, is the deliberate engagement with false information and the manipulation of audiences and information to expressly cause division, harm, and mislead individuals (Puig Larrauri and Morrison 2022: 175-178). Pukallus (2022) addresses weaponised speech under the banner of discursive dehumanisation which is used to destroy civil society's ability to peacefully cooperate, to engage non-violently with difference, and ultimately the social fabric of civil society. For her, discursive dehumanisation draws from an extensive repertoire of basic linguistic tools, which includes the dissemination of fake news, rumours, conspiracy theories, the re-writing of history as well as lies. Discursive dehumanisation often uses metaphors and especially those of disease or pollution, and further engages in objectification. Discursive dehumanisation occurs across the communicative spectrum of civil society - the news media, fiction, in unmediated communication as well as in the arts (both performative and visual) and importantly, divisive content can be spread both offline and online and it is the latter that is of particular concern for DMAPS. Finally, it operates with symbols, scapegoating and stereotyping. Discursive dehumanisation can be used within civil societies against specific groups but it can also be used against other societies, nations or trans-border cultures, identities, religions and ethnicities.<sup>4</sup>

Stereotyping has been identified by DMAPS as a primary issue in sustaining gender inequality and misogyny online, as ideas about women as weak, dependent, obedient and submissive are spread widely and women's challenges to these stereotypes are met with hostility or shut down. For example, participant-led research in Lebanon focused on gender stereotypes and found that patriarchal assumptions about women and gender roles are reflected in social media, with explicit reference to discrimination against women in the workplace, honour crimes, gender-based violence and harassment, and masculinity. The researchers found that Facebook posts that referred to gender or women-related topics were often ridiculed and generated aggressive responses from men, including dehumanising language and religious references to support this position. Participants in in Libya identified similar findings: men tended to post aggressive responses in online discussions about women's rights and freedoms including economic independence, relationships, divorce, and the personal choice to wear the hijab or not. These negative responses were similarly supported by references to religion, with the words 'Muslim', 'women', 'Islam', and 'religion' appearing most commonly and alongside each other. The participants working on OPT found similar topics, reactions, and language used in social media discussions around gender stereotypes and women, but also some in relation to discussions of sexual identities and orientation. It should be noted, however, in the OPT there is less data and the topics of women's rights and gender stereotypes were raised largely by human rights and women's organisations.

<sup>1</sup> DMAPS is also very much based on what in the academic literature is called 'local approaches to peacebuilding/local peacebuilding' and methodologically it uses 'participatory action research' which was mainly developed in the social sciences and born out of a concern to ensure that research is relevant, locallydriven and context-specific. It is very much about the co-production of knowledge. On this see position paper 2.

<sup>2</sup> Although they frame their insights as affordances of technology for conflict, their insights into weaponised communication are applicable and relevant beyond the digital world.

<sup>3</sup> Consortium members of the Universities of Sheffield and Manchester also produced a research background paper on hate speech and polarisation as well as political tribalism for the DMAPS project.

<sup>4</sup> On the relevance of cultural relations see position paper 3.

#### 2. The consequences of weaponised communication

When it comes to the consequences of weaponised communication, theoretical research focuses on questions such as: what does a communicative environment look like during and after weaponisation? On a macro-level, it is fair to say that weaponisation leads to a distorted and polarised communication environment which is characterised by a) a diminishment of relationships between members of society nationally and locally as well as cultural relationships between different societies internationally and b) weakened or destroyed institutions including the news media and the law (Pukallus 2022). On a micro-level it is necessary to have the tools to examine the features and signals of a polarised environment and the specific links that can be drawn to the macro-level of society. Here Build Up have developed five archetypes of polarisation that are vital in the analysis and understanding of polarisation, its manifestations and effects on the micro level of society, i.e. within a specific digital environment (though it could be an offline environment too). The five archetypes are attitudes, affiliation, interaction, interests and norms. Each is further explained with signals of polarisation: signals that indicate polarisation but also and importantly provide a list of criteria that could be used as coding categories for the analysis of any communicative engagement on social media in terms of polarisation (for more information see position paper 2).

More specifically, these five archetypes of polarisation have fed directly into the analysis of social media content and the design of pilot interventions in the DMAPS program. One clear example of how attitudes constitute an archetype of polarisation is the conservative and misogynistic attitudes towards women on social media as found by the research groups working on Libya, Syria, Lebanon, and the OPT. A further example is how individuals' affiliations have been most explicitly manifest in Kurdistan, where hate speech on social media is directed at ethnic and religious minority communities (specifically the Yezidi and Christian minorities). Participants form Kurdistan found that hate speech has been targeted directly at individuals based on their religious affiliation and has included name calling, mocking religious figures and icons, and statements to challenge/undermine (in particular the Christian community's) faith and religious practice. The research also found that whilst the othered communities were discursively attacked online, the affiliation of the Iraqi majority – with Islam – was also referenced in a positive way, presenting it as the 'correct' or 'true' faith and thus its followers as being right and virtuous, further highlighting the ethnic, religious divisions. Interactions, interests, and norms are also invaluable lenses through which we can analyse the polarisation that takes place in all country clusters.<sup>5</sup>

# 3. The peacebuilding potential of communication: building civil capacity and communicative resilience

The third area of research that underpins DMAPS is that which focuses on the solidarising and peacebuilding potential of communication, widely conceived to include the arts. Here DMAPS employs a hybrid approach that includes both locally-led digital and offline interventions to decrease polarisation and conflict. Relevant to this is Puig Larrauri's and Morrison's (2022) work. They have identified digital conflict drivers and respective peacebuilding responses. More specifically, they identify eight possible responses to digital conflict drivers: (1) Platform rules & Terms of Service; (2) Ethical rules & Controls on Platforms; (3) Changes to platform structures; (4) Content moderation; (5) Digital Literacy; (6) Depolarisation & Behaviour Change Programs; (7) Personal and Collective transformation and (8) non-violent communication. These eight proposed responses can be defined as technical/mechanic (1-4) as well as socio-psychological (5-8). The former intervene in the digital technologies directly; the latter are more oriented towards civil society and capacity-building and can (and do) use digital technologies. In this socio-psychological sense, arts and creative initiatives are extremely well suited to the intervention because of their capacity to engage, transform, and allow for expression in diverse ways. In attempts to change behaviours, transform attitudes, and communicate in non-violent ways<sup>6</sup>, artistic and creative practices offer the ideal tools to facilitate this in that, regardless of their form (music, theatre, dance, visual arts, etc.), they encourage individuals to engage with alternative perspectives and express themselves in different ways (Bleiker 2018: 23). For example, scholarship outlines the rehabilitative and behaviour-changing functions

<sup>5</sup> For a deeper dive into findings on the archetypes of polarisation that emerged from participant-led social media analysis please see position paper 2. 6 On the transformative communicative capacity of the arts see Pukallus (2022).

of music, dance and theatre as artistic practices which, when incorporated into (especially community-led) peacebuilding interventions, greatly contribute to social cohesion (see Premaratna and Bleiker 2016: Premaratna 2018: Arthur 2021).

Just as the DMAPS program has particularly aimed to achieve the socio-psychological outcomes (5-8) through the design of its various interventions, artistic practices are ideally matched to these same outcomes and thus the incorporation of arts actors in the intervention supports/complements its aims well. For example, participants in Yemen focused on creating a film, informed by the local community, which aims to address the sectarian and ethnically-charged hate speech that has appeared online. Through the art of film, it is possible to present extremely contentious subjects, deep disagreements and polarising ideas in a more distanced way, and encourage viewers to imagine alternative dialogues and realities to those of their current situation, characterised by division.

Pukallus (2022), in turn, starts with the aim of capacity-building and focuses on the question of what general principles<sup>7</sup> should underpin communicative engagement between former enemies or deeply divided and polarised groups, independently of whether such engagement occurs online or offline and what kind of communication it uses (unmediated, factual and fictional media, visual or performative arts). Her concern was to devise principles that can be locally and contextually interpreted and that can be used as a guarantor for safety as it cannot be assumed that communicative engagement, or attempts at depolarisation or peacebuilding will be non-violent and feel safe for participants. The three principles that Pukallus (2022) has developed are;

Principle 1: Participants have to make a commitment to manage their individual negative emotions (emotional forbearance)

Principle 2: Participants have to make a commitment to both listen to the other and importantly to hear the other (perspective-taking)

Principle 3: Participants have to commit to making only such contributions that are supportive of the pursuit of peace (reasonableness)

These principles of discursive civility act as a guarantor for safety and are therefore defining characteristics of safe spaces both online and offline. Such safe discursive spaces exist across the communicative spectrum of civil society. An arts project (both performative and visual) can be a safe discursive space in three ways: a) the production team of the specific kind of art as a safe space. The team is a safe space when it upholds the three principles of discursive civility; b) the art itself – a dance, a piece of music, a photography has to uphold the three principles of discursive civility in order to be contributing to peaceful cooperation and association and c) the audience when it engages with the art does so in a way that also upholds the three principles. Though these principles might appear abstract, they are indeed intuitively upheld by those who aim to create safe spaces or who create or moderate conversations in our project. For example, the Jordanian participants found that much discussion on Facebook was guite controlled, suggesting a desire on some level to manage emotions. The findings point to attempts to minimise polarisation in some social media circles and even a degree of self-censorship. Various interventions want to bring people closer together through debates, conversations, discussions and new narratives, others will interpret messages through and express them via humour and will have to find a 'safe' tone to do this, and others want to present new visual content. Some have realised the importance of safety they either mention the term 'safe space' or have made arrangements to bring people in that act as moderators for 'ground rules' (discursive civility) - family therapists and moderators for communicative engagement for example. Each intervention that aims to bring polarised groups together has thought about how to make this engagement safe and some have attended a clinic run by Pukallus to help with such needed communicative

7 For more detail on how the principles of discursive civility are upheld in music, memorials, news journalism and theatre see Pukallus (2022, Ch 5).

mechanisms of safety. In short, 'ground rules' upon which non-violent communication relies and that ultimately provide the basis for the development of civil capacity and communicative resilience. These principles also underpin the technical/mechanic (1-4) responses identified by Puig Larrauri and Morrison (2022) and discursive civility also acts as an enabling mechanism for the socio-psychological outcomes (5-8), as well as for the impact the project aims to achieve via its depolarisation strategies. For the latter, discursive civility can act both as an enabling mechanism and a supportive baseline strategy.

# 4. Peacebuilding and arts as communication

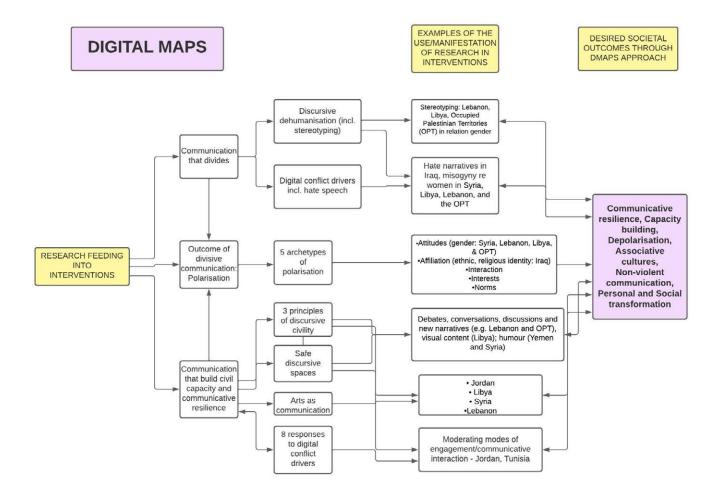
The fourth and final area of research that informs the DMAPS approach builds on the last and sees arts as powerful communicative and influential social tools in contexts of deep division and peacebuilding. Underpinning DMAPS's approach is the idea that not only should peacebuilding be participatory but that arts have a unique role to play in this. Premaratna and Bleiker (2016) argue that arts-based peacebuilding practices are invaluable but that they can only be considered as sustainable if they are undertaken at every level of society, precisely because they require the endorsement and involvement of the communities affected by conflict.

The arts have great significance in sustainable peacebuilding because they can be embedded in, work through, and led by local communities. Art happens at all levels of society and can be manifest in various forms; this renders it appealing across social divisions and classes, making it an excellent medium for fostering connections and relationship-building. Indeed, current research points to three areas to be addressed, which coincide with artistic practices, as noted by Premaratna and Rajkobal (2021): everyday relational engagement through aesthetics, incorporating emotions, and including marginal voices. Further, Arthur (2015; 2016; 2019), who argues that art has a powerful communicative and representative function that gives a public platform to those most under-represented in society. Further, artistic spaces and places where creativity is promoted are important in peacebuilding because they provide an alternative focus, an alternative perspective, and encourage different ways of thinking and seeing the world. Where realities are shaped by violence, prejudice and conflict, this is key to moving forward (Vogel et al. 2020; Bleiker 2018). Such an approach encompasses various arts initiatives produced and/or used for peacebuilding, including theatre, music, dance, cinema, poetry, and sculpture among others. The incorporation of these ideas are evidenced in the planned activities of the Yemeni and Jordanian participants who are engaging cartoonists, graffiti and street artists, political illustrators, social commentators and activists to participate in creative strategies. Further, the principle that artistic practices are inclusive and representative of minority voices is applied in the plans for the Libyan participants; to address polarisation in relation to gender and the marginalisation of women, the group and participants aim to engage with women artists and use their work to instigate conversations about gender stereotypes and gender-based violence.

In addition to its communicative and representative functions, the DMAPS approach also understands arts to exercise influence through their function as social opinion multipliers. Social multipliers indicate the extent to which individuals' behaviour is influenced by social interactions and their peers (Burke 2016). The more an individual interacts socially with peers engaged in a particular activity, the greater the chances of that individual also engaging with the activity. This is then established as a socially normal behaviour, which in turn influences attitudes and norms over time. When applied to arts-based peacebuilding initiatives in safe spaces, these socially normal behaviours could include: creative thinking and alternative imaginings of the world and the future, to see past the violence; creative media of communication and expression; tools for representation and inclusion of marginalised voices; and crucially, the principles of discursive civility including emotion management in order to maintain respectful and constructive dialogue. By engaging an individual in arts-based peace activities as social multipliers, the individual is exposed to others participating in the activity and its associated attitudes. In a context of deep polarisation and division online, to invite individuals to engage in safe conversations about the causes of division and in artistic peacebuilding activities is to introduce them to peace-oriented and inclusive socially normal behaviours. For example, participants in the OPT aim to create conversations and content on social media that tackle gender stereotypes and negative attitudes towards women. The hypothesis is that if gender stereotypes are discussed more openly in the Palestinian virtual space in a safe way, then fruitful conversations which challenge problematic attitudes and behaviours will take place. If these conversations were to gain in popularity, more individuals may wish to participate (in this new socially normal behaviour/dialogue) and escalate from there. Similarly, participants in Syria and Lebanon aim to use humour to create more positive content online about women's rights and gender as a first step in discussions of polarisation in a way that is tailored to the unique context and which is community-led.

# A schematic representation of the DMAPS approach

Before concluding with a summative position statement regarding the approach taken by DMAPS, it is worth summarising the DMAPS approach outlined above and the way in which the theoretical foundations of DMAPS have fed into the pilot interventions in the eight countries of the MENA region in a schematic way as follows:



#### **Conclusion: a position statement**

This position paper shows that the DMAPS program is based on five main propositions which it implements and actively engages with throughout. The first is that theoretical research – both academic and practitionerproduced provides a vital source of knowledge that can be used to inform and shape practical approaches and interventions. In the case of DMAPS, we have shown the importance of the research related to social media polarisation, digital conflict drivers and weaponised communication, to the role of communication (including the arts and digital tools) in peacebuilding within and across civil societies and the way it can inform approaches that aim to (re-)build the civil capacity for peaceful cooperation, communicative resilience as well as social relations. Second and directly related to this is the recognition that civil societies – whether solidarised and cooperative or divided and polarised – are deeply communicative in nature. This is a recognition that underpins the entire DMAPS program. Third and contained within the previous point yet worth stating separately, the dangers of online polarisation through social media are real and social media has the potential to affect the capacity for peaceful cooperation offline. This is why DMAPS recognises digital peacebuilding as a branch of communicative peacebuilding as an urgent imperative and vital to social relations both offline and online. Following on from this and fourth, the DMAPS program recognises that peacebuilding requires a hybrid approach - offline and online (digitally). This has been concretely shown in pilot interventions in which participants created offline communicative outputs and initiatives that were then shown on social media. Here social media acts as dissemination platform to reach certain audiences thereby harnessing its scale and spread for peacebuilding aims. What this means is that communication in society is inherently hybrid and that this hybridity needs to be reflected in peacebuilding as well.<sup>8</sup> Finally, fifth, and crucial to DMAPS is the need for participatory action research and local approaches to peacebuilding. It is the only way in which local problems can be deeply understood, researched and effective and relevant (hybrid) peacebuilding interventions can be designed. DMAPS is participant-led and in this phase involved 18 local partners.

# References

Arthur, C. (2015) Writing national identity on the wall: The geração foun, street art and language choices in Timor-Leste. *Cadernos de Arte e Antropologia*. 4(1): 41-63.

- (2016) Painting their past: the Geracao Foun, street art and representing notions of East Timorese-ness. *Sojourn: Journal of Social Issues in Southeast Asia*, 31(1): 173-206.
- (2019) Political symbols and national identity in Timor-Leste. Basingstoke: Palgrave Macmillan.
- (2021) Street Art in East Timor: Creative (Re) Constructions of Identity in Times of Crisis. In:

Campos, R., Pavoni, A. and Zaimakis, Y. (eds) *Political Graffiti in Critical Times: The Aesthetics of Street Politics.* New York: Berghahn Books, pp. 275-301.

Bleiker, R. (2018) Mapping global visual politics. In: Bleiker, R. (ed.) *Visual global politics.* London: Routledge, pp. 1-29.

Burke, M. (2016) Social Multipliers. In: Vernengo, M. et al. (eds) *The New Palgrave Dictionary of Economics*. Basingstoke: Palgrave Macmillan, pp. 132-155.

Locally Driven Peacebuilding (2015) Signed letter, 27 March, available at: https://www.cla.purdue.edu/ppp/ documents/publications/Locally.pdf

Premaratna, N. (2018) Theatre for Peacebuilding Basingstoke: Palgrave Macmillan.

Premaratna, N. Bleiker, R. (2016) *Arts and theatre for peacebuilding*. Richmond, O., Pogodda, S. and Ramovic, J. (eds.) *The Palgrave Handbook of Disciplinary and Regional Approaches to Peace*. Basingstoke: Palgrave Macmillan, pp. 82-94.

Premaratna, N. and Rajkobal, P. (2021) *The arts as a peacebuilding approach: Feminist contributions and directions.* In: Väyrynen, T. et al. (eds.) *Routledge Handbook of Feminist Peace Research*. London: Routledge, pp. 260-269.

Puig Larrauri, H. and Morrison, M. (2022) Understanding digital conflict drivers. In: Mahmoudi, H., Allen, M. and Seanan, K. (eds.) *Fundamental Challenges to Global Peace and Security. The future of humanity*. Basingstoke: Palgrave Macmillan, pp. 169-200.

Pukallus, S. (2022) *Communication in Peacebuilding.* Civil Wars, Safe Spaces and Civility. Basingstoke: Palgrave Macmillan.

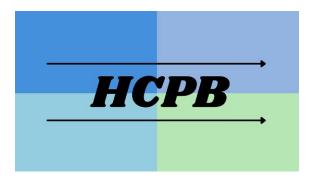
Vogel, B. et al (2020) Reading socio-political and spatial dynamics through graffiti in conflict-affected societies. *Third World Quarterly* 41(12): 2148-2168.







The University Of Sheffield.





The University of Manchester





Produced by Build Up for the DMAPS program

Participatory Action Research, Polarisation, and Social Media:

Ongoing lessons from the Digital MAPS program



Digital Media Arts for an inclusive Public Sphere (Digital MAPS) is a program that supports young leaders from creative and media arts organisations across eight countries to conduct social media mapping in order to understand how issues of identity, social cohesion and inclusion manifest on and are shaped by social media. This participatory action research is then used by these organisations to design pilot interventions that counter polarization and promote inclusivity and openness in the networked public sphere. The program is implemented by Build Up, Sheffield University, Manchester University, the American University of Kurdistan and datavaluepeople, in collaboration with the British Council.

The DMAPS program is of interest to both practitioners in the peacebuilding and creative arts spaces and to academics researching how conflict dynamics intersect with online space and communication practices because it pilots a distinct approach to affective polarisation. The program is premised on a recognition that affective polarisation is something that is happening to users of social media – it is built into the very logic of the social media platforms – and is arguably exacerbating existing fault lines. Thus, the emerging norms of digital culture and their effect on interests, attitudes, affiliations and interactions of social media users present a risk to the type of inclusive public sphere required to build strong open and democratic cultures. This framing is an important shift from mainstream interventions in the network public sphere that focus on content moderation (including identifying and taking down hate speech) and digital literacy (understood as identifying and avoiding misinformation), missing the full complexity of how affective polarisation operates on social media.

In order to share the thinking that has informed this innovative approach and our evolving learnings from this program, we are producing two papers for discussion. This paper outlines the participatory action research process and provides examples of how the resulting social media analysis identified signals of polarisation. It also briefly discusses challenges and lessons from this practitioner-led social media analysis process. The other paper outlines the four areas of theoretical research that support the program, and how these translate into a research-design journey for participants.

# Participatory action research applied to social media mapping

The DMAPS program applies participatory action research to social media mapping in order to explore the details of polarizing dynamics that unfold on social media. The participatory action research process comprises four participant-led steps:

# 1. Defining problem statements

- 2. Collecting data
- 3. Coding data
- 4. Analysing and designing action

Throughout the four steps, participants are in the lead, making this a distinct approach that ensures research findings are contextually grounded and can directly connect to the design of relevant actions.

# 1. Defining problem statements

Participants start by formulating problem statements that define the scope of the social media mapping. Problem statements comprise a conflict / polarisation thematic area and a series of research questions, and list the country (or countries) and social media platform(s) of interest. Partners worked within regional clusters in a facilitated process to identify their most pressing issue areas and research questions. For example, in Yemen the problem statement focused on expressions of religious and ethnic identities on Facebook and Twitter, addressing the following research questions:

What identities are Yemenis expressing online?
What terms and expressions do Yemenis use to describe their identity online?
What other topics are discussed along with identity?
Do conflict fault lines show up in the expression of identities online? If so, how?
What actors make statements about Yemenis? Are these actors "grassroots" or in positions of power?
Are there identifiable networks of identity on social media?
How do social media accounts cluster into networks using similar terms and expressions?
Do people who express different identities interact with each other? If so, how?
Do polarising conversations about identity attract a certain type of actors?

# 2. Collecting data

In order to collect data relevant to their problem statements, teams decide on what concrete sources of social media data they want to monitor, producing a detailed list of accounts (public Facebook pages & groups, Twitter handles, YouTube channels) and keywords (to search on public posts on Facebook, tweets or public YouTube videos). These lists are a way of narrowing the social media sphere to a scope that is deemed relevant to the problem statement. Data from these lists was then scraped (i.e. automatically collected), with new data added to a database every three days.

# 3.Coding data

Data collected (scraped) comes with some useful information about the date it was posted, who posted it, and how many reactions it had. However, with thousands of pieces of data, the research process required additional data coding to organise the information. Classification models were built for topics, actors and sentiment.

Topic labels indicate the general topic (or sub-topic), such as "arrests" or "economic collapse" or "covid / vaccination". Topics might also include incidents or behaviors related to the conflict / polarisation theme identified in the problem statement, such as "hate crime" or "self-censorship" or "harassment" or "kidnapping". Topic tagging was also used to group posts / tweets / videos / comments into relevant and irrelevant groups: those with a topic label are relevant since they are about a topic of relevance to the problem statement, while those with no topic label are irrelevant. Topic labelling starts with a manual labelling of approximately 2000 posts / tweets / videos, where team members looked at the text to assign a label, and then identified what features of the text (words and phrases) indicate that this label should be applied. This logic is then repeated by an automated model and applied to all other data collected.

Actor labels indicate what group(s) the author of a post, tweet or video belongs to. Actor labels might include affiliation to a group that is a party to the conflict, country where they are located, profession, or type of audience they attract. Actor labels were manually applied to the data source list and then populated automatically to all data collected.

Sentiment labels indicated the semantic sentiment, that is whether the language used in a post / tweet / video is positive, negative or neutral. Sentiment labels were automatically applied using AWS comprehend, a trained sentiment analysis model.

# 4. Analysing and designing action

Once the data coding was finalized, all the data collected was presented in a dashboard designed to support participant-led analysis. The dashboard comprises tables and graphs that support topic, actor and sentiment analysis. For topic analysis, graphs and tables show the main topics discussed, who discusses them, what words are most commonly used to discuss it, and what topics are discussed together. For actor analysis, graphs and tables show the main actors and how they are connected in the network public sphere (where two actors are connected if they comment on or retweet the same content). For sentiment analysis, graphs and tables show what sentiment is most common for any topic or actor, using both semantic sentiment labels and (for Facebook) emojis. Finally, the dashboard makes the entire dataset searchable, both

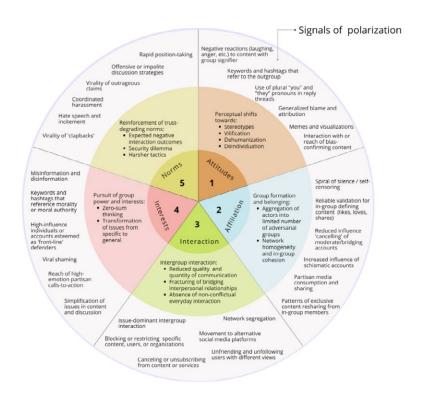
using a series of filters (actor, topic, sentiment, date) and through a keyword search (applied to the text or account name).

For topics, actors and sentiment, the dashboard enables participants to conduct quantitative, qualitative and network analysis. Quantitative analysis, conducted using the bar and pie charts provided, tells us what people are talking most about, who is talking about this, and how are people reacting to it. Qualitative analysis, conducted using the tables provided to search and read content, tells us how people are talking, and what the tone and emotion are in a narrative. Network analysis, conducted using the interactive network graphs provided, tells us who is talking about what with whom, what voices and topics are on the margins, and which ones are bridges. Overall, the dashboards allow participants to not only answer their specific research questions, but also to come up with key findings that outline what the data tells us about what conversations selected actors are having on social media, how they are connected and where participants' own social media accounts are placed within these conversations.

# Identifying signals of polarisation

This paper does not summarise all the findings that each team arrived at, nor the general finds for each country. Instead, it aims to provide an overview of the type of findings that participants arrived at, in order to offer an understanding of the opportunities of participatory action research applied to social media mapping. Concretely, this section provides examples of how the resulting social media analysis identified signals of polarisation.

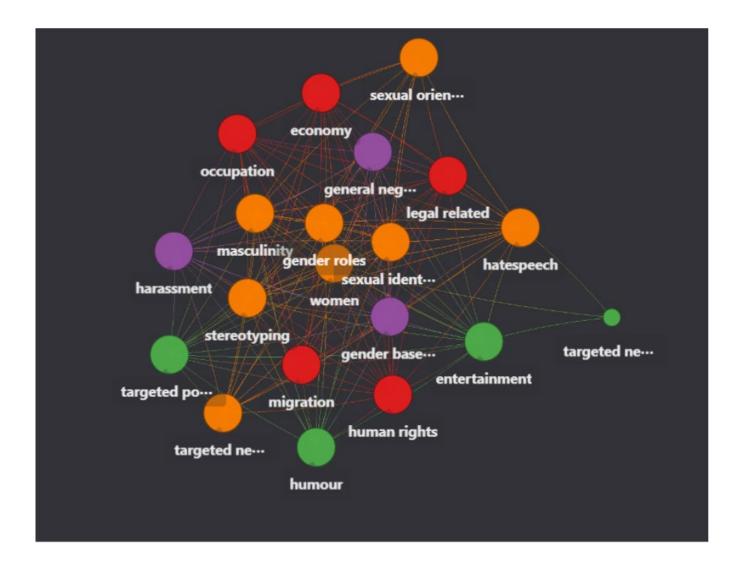
The Digital MAPS approach recognises that the attention-for-profit platform logic of social media leads to segregated networks and spreadable spectacle, with emotive content rewarded by algorithms and therefore more likely to be viral. To concretise this into an approach to analysis, participants were introduced to a framework for problem analysis that specifies the roles of social media as a digital conflict driver that powers affective polarization through influencing identity construction, incentives, and discourse, and maps the induced behaviors of social media users into observable archetypes of polarization within the network public sphere.



In conducting the social media mapping, participants found concrete signals in their data that pointed to each of these archetypes of polarization. Where there are other ways to organise the findings of the social media mapping process, we find that this categorization points most directly to how research findings eventually informed the design of pilot projects.

# **Attitudes & Norms**

Participants from Syria found evidence of polarizing attitudes on social media posts discussing the roles women can play in society, where there is clear evidence of a perceptual shift towards stereotypes. This is clear not just in a qualitative analysis of posts, but also in observing the topic network graph, which shows that the issue of the role of women is largely addressed in terms of masculinity and gender identity. Digging into these posts showed that conversations revolved around the appropriate jobs for women based on physical condition or customs and traditions. On the other hand, the role of women is not often discussed in relation to human rights or (even less) economic issues including material needs.



Looking at the same data, participants in Lebanon and Palestine observed a polarization of attitudes towards women that has resulted in the perpetuation of polarized norms in the network public sphere. Men have a greater tendency to stereotype women, as was clear in qualitative analysis showing frequent comments that suggest emotions are a sign of weakness in women, or sentiment analysis showing negative reactions towards posts indicating that women have financial and educational independence. On posts addressing issues of gender on Facebook, women use mostly love reactions where men use angry and laughing reactions, suggesting that men tend to ridicule this content. The wordclouds show a pattern where the word "woman" or **ödia** is always joined with a male figure jack. خيش جوز It was evident from wordclouds and qualitative analysis that men are commenting on topics predominantly about women (such as economic independence, custodial rights, marriage and divorce, gender-based violence) all the time, whereas women tend to be sidelined when it comes to topics predominantly about men. In general, participants observed that men either react negatively to content about women (quoting religious ideas or arguments) or approach it with ridicule (stereotyping women by using humor).

The analysis results on polarised attitudes and norms around gender in Lebanon, Syria and Palestine informed the content strategies of three teams. Participants in Syria are running a campaign to challenge stereotypes about women in scientific positions. Participants in Palestine have set out to influence a change in the discourse and attitudes towards gender roles and stereotypes by producing content that tackles these issues and that calls for action and discussion of such topics, more specifically, women assuming certain jobs and positions. Participants in Lebanon have designed an approach that uses humor and gender conscious content to mainstream gender equality and break gender stereotypes.

# **Attitudes & Interests**

In Kurdistan, a qualitative analysis of language used on Facebook posts revealed a toxic and polarizing atmosphere against religious minorities (especially Yazidis and Christians), fueled by misinformation and dehumanization, signaling both attitude and interest polarization. This was contrary to the initial hypothesis of participants that hateful narratives would center around ethnic identity. Posts that are related to religious ceremonies (especially during religious festivities) draw discriminatory comments and mockery against religions symbols. In addition to discriminatory comments, Islamic verses and sayings of Prophet Mohammad are posted in the comment sections; these verses are not inherently problematic, but are being instrumentalised to support hateful narratives. The network of Facebook commenters shows that the commenters are normal individuals with their original identities, with no remarkable connection that could be drawn between the accounts. A unifying characteristic of comments is the use of plural "we" and "you" – in reference to superiority. As a result of insights on attitude and interest polarization, participants in Kurdistan are changing the strategy on their own page, by creating a dedicated group of moderators who intervene in comment threads where misinformation and dehumanising narratives towards religious minorities is shared.

In Jordan, the language used across Facebook and Twitter shows signs of the use of morality and moral authority to polarize interests. Participants observed that social media users use words that show moral anger, such as (shame, taboo, homosexuality) and were more likely to respond to emotional words about religion (such as heaven, hell, haram, wrath of God). They also noted the influence of religious influencers on people's opinions. For example, the Convention on the Elimination of All Forms of Discrimination against Women (CEDAW) was published by Dima Tahboub, who posted her opinion that some of the provisions of the agreement contradict the Islamic religion, and attracted many comments repeating this moral stance. More generally, mainstream media pages often publish questions about other religions or share videos of people speaking against other religions. Where these posts are disguised as open exchanges, they attribute moral authority to one position, attracting comments only from those who agree, and often leading to hateful or violent comments.

Identifying this tactic to fuel polarization informed the content strategy of program interventions by two participating teams. One team is intervening to raise awareness of the importance of creating safe comments on Facebook, especially those related to the posts of media actors and religious actors that drive polarization. Another team noticed that commenters on social media invoke freedom of opinion and expression to justify hatespeech and incitement to violence, and is targeting content addressing this issue in relation to the topics where comments most fueled polarization, namely religion, customs, traditions, and women's issues.

### Norms

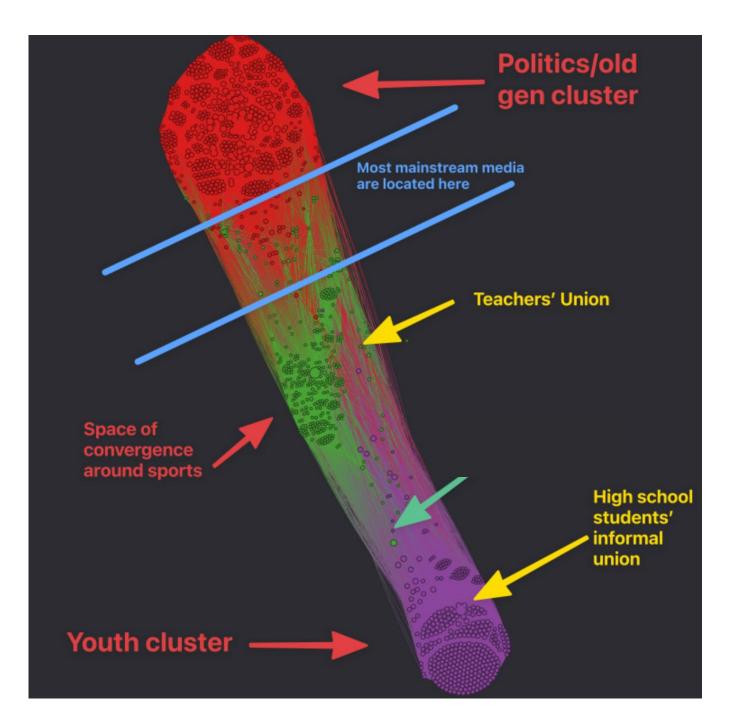
In Yemen, the prevalence of coordinated harassment and back-attacks (exclusion from the group based on identity) on social media was an evident signal of polarised norms. In the Yemeni retweet network, there are three clear poles representing the three conflicting political parties. These relatively homogeneous subnetworks present discourse directed exclusively to the same group, with repeated patterns of content sharing by the same group, including lots of negative comments and hate speech on accounts, especially on posts that talk about the group. The rate of adoption of hate speech on Twitter is high, and that its actors are individuals / influences and not external platforms or institutions.

Identifying the pervasiveness of hate speech resulted in all three participating teams from Yemen focusing on this. One team is running a narrative change campaign to counter hate speech by offering a view that highlights the peaceful coexistence between North and South before the conflict started. Another team is running an awareness raising campaign on hate speech and its impact on society, while a third team is focusing its content on sharing people's experience of hate speech.

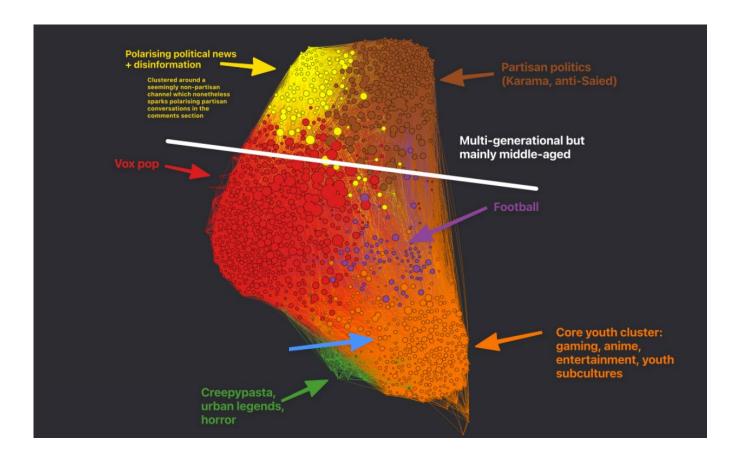
# **Affiliation & Interactions**

The Libyan social media map (focused on conversations about gender) and the Tunisia social media map (focused on understanding differences between older and younger generations) both provide evidence of the difference between polarized affiliation (where aggregation of actors, homogeneity, self-censorship, echo chambers) and polarized interactions (where network segregation, limited bridging).

The Facebook commenter network for monitored accounts in Tunisia shows clear signs of network segregation (a signal of polarized interactions) with some bridging (a possible space for intervention). There is a clear boundary between a youth cluster (with content mostly from alternative media and activist accounts) and a cluster dominated by the older generation (with most content coming from mainstream media and partisan political accounts). The most influential account in the youth cluster is such as Lyceena, a page which represents the voice of high school students and is engaged in a struggle against the education establishment which is dominated by older people. There is a space of convergence around football (centred on Foot24), with commenters on this content bridging between the two poles.



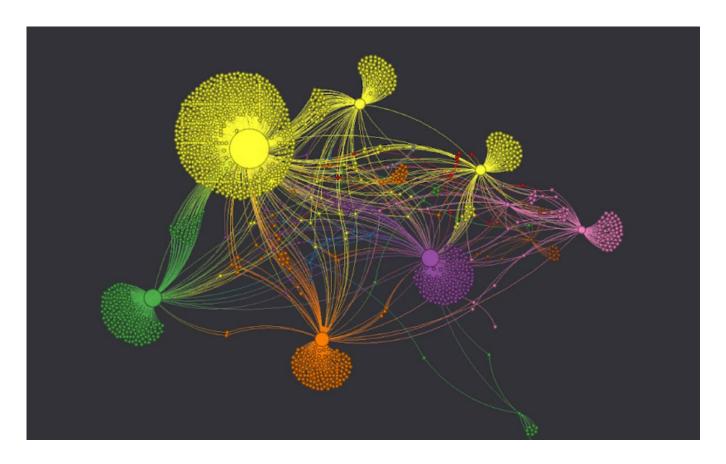
The YouTube commenter network of monitored accounts in Tunisia shows signs of an aggregation of homogenous actors into distinct clusters, a signal of a network that is likely to be dominated by a polarised affiliation. On one end of the graph, there is a youth subcultures cluster which is dominated by entertainment. On the other end of the graph, there are two political clusters that tend to attract middle-aged people whose content is highly polarising and often spreads disinformation. Between them, there is a large cluster around vox pop content which suggests a strong appetite for content which features the voices and perspectives of ordinary people.



The analysis results on polarised affiliation in Tunisia have led participants to design an intervention that centers the voices of young and old people sharing their perspectives in a video to start a conversation, in an attempt to approach the "vox pop" style that appears to act as a bridge between the youth and older generation cluster.

The Facebook commenter network of accounts monitored in Libya shows clear signs of network segregations with limited bridging, signals of polarized interactions. Overall, most accounts monitored are not connected regularly by commenters, including all of the participants' accounts. For those that are connected, there is a small pole (orange) made up exclusively of Jaafar Tok DW and his tight network of commenters, who are not engaged in the rest of the conversation.

DW is the account on Facebook with the most engagement, and he deliberately creates a space for controversial posts in the way he writes captions, he invites people to disagree in the comments section and encourages engagement. There are very few bridging accounts who comment on both DW and pages in the other pole (six concrete accounts were identified).



The network of retweets by accounts monitored in Libya shows clear signs of an aggregation of homogenous actors, a signal of a network that is likely to be an echo chamber dominated by a polarised affiliation. All the main nodes are activists working on gender. Interestingly, one of the participant accounts sits in the centre of this echo chamber.

The analysis of polarised aggregation and affiliation in Libya has defined the targeting strategy used by some participants, whose focus is now to bridge the gap between its circle of influence on social media platforms and where the major discussions on gender roles are happening.

# Challenges & lessons from practitioner-led social media analysis process, and a position for future work in this area

Two key challenges have emerged in the participatory action research process described above. First, the research only looks at a sub-space within the broader network public sphere, and analysis is therefore limited by the problem statements defined by participants and by selection bias in relation to data sources. Second, it takes time to set up the data collection and coding needed to perform this kind of in-depth analysis, and to acquire relevant social media analysis skills. Some participants more focused on intervention / activism do not see the value in performing participatory action research and would prefer an external research report that provides evidence and supporting inputs to their programming.

These challenges point to the main lesson learned from the analysis process of the Digital MAPS program: that there is a clear opportunity for conducting research that informs evidence-based interventions that address different aspects of polarization in the network public sphere, but sustained participatory action research may only be worthwhile for actors who expect a sustained engagement in digital interventions to shift polarization in

the network public sphere. Actors who would want a one-off input may benefit most from externally produced research on a snapshot of the network public sphere, rather than an ongoing participatory analysis process.

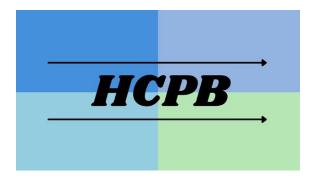
Despite these challenges, the examples above illustrate that the participatory action research process resulted in participants identifying concrete signals of polarization in a sub-space of the network public sphere relevant to their problem statements. The clearest opportunity arising from this process is that these insights, because they are concrete, can be directly translated into program design. Concretely, participants used research insights to define content production, change targeting strategies, and refocus their own social media activity. Perhaps most interestingly, the analysis process in some contexts resulted in participating teams designing interventions that used different tactics to address the same identified signals of polarization, which makes a shift in the network public sphere as a whole more likely. We believe this pilot program therefore demonstrates the value of participatory action research on social media, and its potential to inform the design of interventions to address polarisation.







The University Of Sheffield.





The University of Manchester





Produced by Build Up for the DMAPS program

Iraq Research Report

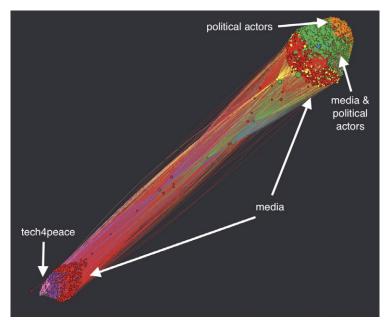


This research report is based on analysis of 197 Facebook pages and groups from Iraq and Kurdistan. These pages and groups were chosen by the three Digital MAPS program partners in Iraq and Kurdistan as the places on social media where they believed most hateful narratives against ethnic identities take place. Posts and comments collected from these pages between November 2021 and March 2022 were labelled by topic (based on the text) and actor (based on who posted). The list of 37 topics and 8 actor types developed by the program partners is available <u>here</u>.

This report outlines insights about actors and topics based on content from these 197 pages and groups. Over the period of November 2021 to March 2022, we reviewed a total of 61,152 relevant<sup>1</sup> Facebook posts in Arabic and 34,671 relevant posts in Kurdish. In total, the 61,152 posts in Arabic gathered a total of a little less than 2 million comments (of which this report analyses 65,259 in detail) and the posts in Kurdish gathered 11,815,066 comments (of which this report analyses 39,074 in detail). Where this report is based on a comprehensive analysis based on the selected data, it is important to remember that these insights do not reflect the entire social media landscape in Iraq and Kurdistan. The final section on 'Archetypes of Polarisation' includes recommendations on possible programming responses.

# **ACTOR MAPPING**

The network of people commenting on Facebook posts reveals two distinct ecosystems (Arabic pages and Kurdish pages), with some bridges. The graph below shows there is a clear division between those commenting on Kurdish (top right) and Arabic (bottom left) pages, and also shows there are a number of pages and commenters who connect these two poles (mostly Kurdish-speaking). Concretely, media pages act as a more effective bridge between the two. In the Kurdish Facebook sphere, Kurdistan24 is well positioned as a bridging page. People commenting on political actors in the Kurdish Facebook sphere are least likely to connect with people commenting in the Arabic Facebook sphere. People who comment on the Tech4Peace page do not comment in the Kurdish Facebook sphere. Kurdish users generally speak Arabic and Kurdish, whereas Arabic speakers don't speak Kurdish, which further validates that some Kurdish pages can act as bridge builders. In general, the



Facebook commenter network (Arabic & Kurdish)

1 Posts of relevance in this report refer to posts that are labelled under one or more topics.

Kurdish pages monitored had a much higher rate of interaction than the Arabic pages. For example, the post with most interactions in Arabic gathered 303k interactions, whereas its equivalent in Kurdish gathered 1.2 million interactions.

Looking more closely at each of these poles reveals where polarisation is located within each one. A small number of pages in Arabic (7) and Kurdish (6) are responsible for the majority of posts made in this period (and this ranking remains largely unchanged month by month). The top 7 pages posting in Arabic were:

- Al Rashid (قيئ الحيش رل المانة)
- 2. Al Farja (قجرف ا قانق)
- 3. Al Rabiaa TV
- I News (قيىئاض ف ل أوين يأ قان ف),
- 5. Kurdistan Rebaza barzaniye Nemir
- 6. Al Khuwa Al nadifa
- 7. Al falouja

The top 6 pages posting in Kurdish were:

- 1. Xendan
- 2. Kurdsat news
- 3. Kurdistan24
- 4. Rachlaken news
- 5. Rudaw
- 6. Xani agency

All the top pages by number of posts were media pages sharing news and updates. We found that many of these posts were likely to be receiving polarising reactions. In the Kurdish Facebook sphere, media accounts in general triggered the most 'angry' emojis. Interestingly, in the Arabic Facebook sphere, Kurdistan24 received the most 'angry' emojis in the Arabic digital space (13.64% of all their posts). In general, it is possible that the more diversified base of followers that media pages usually get, especially when their political affiliations are not clearly stated or publicly shared, can explain angry reactions to many posts.

In the Arabic Facebook sphere, we found that the pages with the most comments, although not the same as those with the most posts, were also all media pages. AlShariqya News (which is not a top poster) gathered the most comments (4 million). They also authored the most shared post (10.4k shares) in Arabic, which was a speech by a right wing shite political leader of the Peace Companies Moqtada Al Sadr. The other pages with the most comments were Al Rashid (top poster), Al Khuwa Al Nadifa (top poster), Al Rabiaa TV (top poster) and Dajla channel (not a top poster).

There are 3 specific pages (AlSharqiya news, Al Khowa Al Nadhifa and Steven Nabi) that both get a lot of comments on their posts and share the most posts with negative sentiment. A qualitative exploration of their posts shows that the news they share isn't necessarily misinformation, but rather is framed in a way to distract the reader from the information they are sharing, for example by mentioning the name or adding a picture of a controversial political or religious leader to a post talking about football. The comments section on posts is then dominated by this controversial association (e.g. the political or religious leader), and reveals that commenters become heavily polarized.

Al Khuwa Al Nadifa page merits special attention because they authored three of the five top posts with the most comments. Local actors have argued that the page is paid by political groups to spread news in a specific way, and these top three posts reveal this approach. One post (with 260k comments) is about the Iraqi security forces. Specifically, the post reminds readers that there are two days left before the launch of the most violent campaign against motorcycles. The choice of words in this caption is biased and not just intended to inform; the caption reads 'Two days left before the launch of the most violent campaign against motorcycles and will be launched on the Iraqi Police's National day'. Another post is about Iraqi elections and sectarian identity (shiite), and gathered 225k comments. The post's caption details that the Parliamentarians of the Sadrist Bloc have entered the 50th Parliamentary mandate covered with shrouds and

carrying rose water. Another post also about Iraqi elections and identity (shiite) that gathered 179k comments reads: 'AI Sadr tweets: A Government with a majoritarian Iraqi decision making has no place for militias and the law will be elevated by an honest Iraqi judicial body'.

Unlike in the Arabic Facebook sphere, on the Kurdish Facebook sphere the pages with the most comments are not exclusively media pages, and also include political actors and public institutions. From the list of top 10 posts with most interactions, 5 are shared by Xendan (media), 3 by Rudaw (media), and 2 by Kurdistan CT (public institution). When looking at the pages that gathered the most comments, we found that Rudaw (the fifth top poster) gathered the biggest number of comments with 3million comments followed by Xendan (2.6 million comments), Kurdistan24 (1.5 million comments), Masrour Barzani (political actor, 600k comments) and Wishe. net (581k comments).

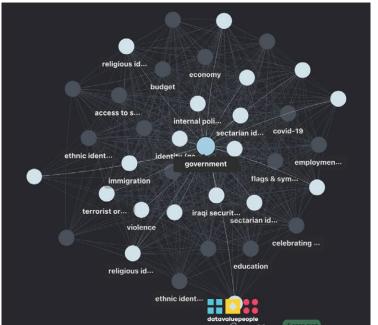
Political actors deserve a closer look. The network graph above shows that, political actors in the Kurdish Facebook sphere generate engagement among a network of commenters; this is not the case in the Arabic Facebook sphere. However, political actors were the actor type that received the highest proportion of 'love' emoji reactions (over 50% of all their posts), suggesting most people engaging with these posts are likely to be unequivocal supporters. This is further corroborated by sentiment analysis for Arabic posts: political actors had the biggest share of 'positive' sentiment-posts. This is valid given the nature of politics and how it is used online to show mostly the progressiveness or tolerance of political actors, particularly in Iraq today. Interestingly, the page that had the most negative sentiment in the entire period of this research is Tech4Peace\Listen, where 25% of posts on the page generate a negative sentiment. This page was created with the intention of reducing polarisation by exposing followers to awareness raising information about the consequences of hate speech online.

# **TOPIC ANALYSIS**

Although posts in Arabic and Kurdish address different topics, in both Facebook spheres users are often exposed to content related to identities. In Arabic, the top 5 topics were government (15k, almost 25%), identity general, Iraqi elections, sectarian identity (shiite) and ethnic identity (Kurdish). It's important to note that, taken together, posts about the majority groups (shiite and Kurdish) are the most common topics of discussion. In Kurdish, the top 5 topics were ethinic identity (Kurdish), (15k, almost 50%), foreign countries, immigration, identity general and violence.

It is interesting to note that where Arabic Facebook posts often discuss Kurdish ethnic identity, the same is not true of Kurdish Facebook posts discussing shiite identity. Whereas there are hardly any posts with identifiable hatespeech in the Arabic posts, hatespeech was the seventh most common topic in Kurdish Facebook posts, with just under 3000 posts. Furthermore, the topic 'celebrating arts & culture' appeared every month in Kurdish in the top 5 topics discussed, whereas it did not appear at all in the top 5 lists of any month in Arabic. This validates our impression from a qualitative review of posts that the Arabic Facebook sphere has more ethnic-centred (potentially polarising) conversations, whereas in the Kurdish sphere conversations about Kurdish ethnic identity are connected to a celebration of arts & culture (and not juxtaposed to another (shiite) identity. Most of the main topics mentioned above - (government, identity general, Iraqi elections, sectarian identity (shiite), and ethnic identity (Kurdish) - are all closely connected and centred in the network of topics, suggesting these five topics appear in most conversations on Facebook (often being discussed alongside a variety of other topics). These connections between topics are stable across the five months observed in the data.

Four of the topics with the most comments - violence, immigration, ethnic identity (arab) and foreign countries - are closely interconnected, but less central to the network of topics, suggesting there may be more engaging posts that only discuss these topics without connecting to other topics. In both Arabic and Kurdish, violence is the topic with most comments, closely followed by immigration. The top posts in both languages are either about violence, or immigration, or both. This high engagement rate and co-related topics suggest there may be more polarising conversations in these comment threads (and indeed the topic of hatespeech also appears close to these four topics).





There are two important differences between Arabic and Kurdish in terms of engagement. First, Arabic posts about right wing shiite political parties and more generally shiite identity gather many comments (and hardly none in the Kurdish sphere). This suggests the issue of shiite identity is either not of interest or not polarising among Kurdish Facebook users. Second, across the board, interactions in Kurdish are much higher. This could be the result of a selection bias (with respect to the Kurdish sources monitored) or may reflect an overall more active Facebook conversation in Kurdish.

Sentiment analysis is of limited use in this analysis. First, there is no sentiment analysis for the Kurdish posts, so a comparison with Arabic posts is not possible. Second, 83% of Arabic posts are classed as having "neutral" sentiment, likely because they are news items (and the negative or positive sentiment is in the comment threads). That said, neutrality can sometimes be used in the Iraqi context to ensure the safety of the poster, especially with the crackdown on freedom of opinion and speech online.

More interesting is a comparison of emoji reactions to posts (see table below). Overall, this comparison further supports the analysis above that the Arabic Facebook sphere has more ethnic-centred (potentially polarising) conversations.

Top 5 Topics with most 'Angry' emoji

# Kurdish

Foreign countries Identity (general) Violence Immigration Ethnic identity (Kurdish)

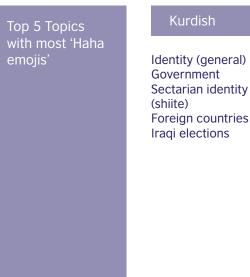
# Arabic

Identity (general) Violence Government Sectarian identity (shiite) Iraqi security forces

#### Notes

This comparison shows that in the Arabic Facebook sphere, users are angrier about issues related to internal politics and internal governance, than they are about external issues (like foreign countries). Whereas in the Kurdish Facebook sphere, users are angrier about 'foreign countries' than they are about sectarian or ethnic identities.

#### **IRAQ RESEARCH REPORT**



Sectarian identity (shiite) Government Iragi elections Identity (general) Foreign countries

Arabic

# Notes

A mix of anger emojis and haha emojis can be a signal of polarisation on a specific post and one that merits deeper analysis. Sectarian identity shiite is the topic that generated the most 'haha' in the Arabic Facebook sphere, while it is also the 4th topic with most 'angry' reactions. Interestingly, 'iraqi elections' did not show up in the top 5 topics with most angry emojis, but it is in the top 5 topics with most 'haha', suggesting that users are more sarcastic  $\$ dismissive when reading about the elections than they are angry. In the Kurdish Facebook sphere however, the topic government did not gather as much anger emojis as it did 'haha' emojis, suggesting that this topic is also responded to more sarcastically.

#### **TIMELINE ANALYSIS**

There were four major peaks in posts on Facebook in Arabic and another 4 major peaks in Kurdish. The most active conversations in these peaks further confirm the trends identified in the sections above.

In Arabic, these four peaks were:

## December 2nd 2021 with 556 posts

December 2021 in Iraq had a lot of events, mainly centring on the results of the parliamentary elections (3) months earlier) that had still not resulted in a government being formed. On December 2nd, a big part of the conversation was about statements and declarations from different political parties and politicians who are either speaking against what they claim to be corrupt elections or show support for the results. Most of the posts shared on this day were statements from AI Sadr or AI Sadr party following a meeting they had on that day. The post that gathered the most interactions on that day is a post from Al Sadr about a confrontation in the metro, which the post captions as "a conflict that happened between a British man and a Muslim man, because the Muslim man was reading the Quran in the metro", showing how the choice of words from media agencies contributes to further divisions). Another post that gathered a lot of polarising interactions is from Steven Nabil who shared a piece of news about a natural disaster that happened in a specific place in the world and the caption reads: 'God all mighty loves Iraq and has chosen to put it in a zone safe from natural disasters'.

#### January 9th, 2022 with 822 posts

Following the appeals and Supreme Court's approval of the election results, on this day in Iraq, the headlines were about the forming of the new parliament. This included a series of news pieces about different statements from different politicians, the voting on the head of the parliament and results thereof, etc. The post that had the biggest number of interactions on this day in Iraq (Arabic) is a post by Al Khuwa Al Nadifa with pictures of the new members of parliament dressed with white textile covering their bodies (this is the textile that deceased people are wrapped in) and carrying rose water. The reason why this created such polarisation in Iraq is because the political meaning of wearing these to a parliamentary session is the following: 'we are the followers of Mohammad Al Sadr (who was assassinated and was under threat when Saddam Hussein was the head of the country) and we follow his footsteps (to mean that they do not fear death and are ready to die for the cause). The caption in itself is neutral, but the reactions on the post are very polarised most possibly because of the nature of the elections that happened in November 2021 where the people were divided between believe the

#### results or appeal.

### January 31st, 2022 with 521 posts

The post that got the highest interaction on this day in Iraq (Arabic) was one about a groom who was surprised to see 'his' bride replaced by another woman on the day of the wedding. However, the reason why there was a peak on this day was because of the forming of the government with statements issued from all sides calling on negotiations to take place and asking for the immediate stopping of terror and violence, in addition to a number of posts reporting on meetings between different coalitions and Al Sadr.

#### March 26, 2022 with 641 posts

A big majority of the posts on that day are about internal politics, from statements from political leaders, to statements from Al Sadr specifically, to updates about the situation in Afghanistan with one specific post that seemed to gather a large number of interactions (6th relevant post on that day with the highest interactions) is a post that is supposed to be about football fans heading to a match, however, the caption reads: 'Among them the marty Omar Saadoun... 3 thousand fans head to [..]', according to our local partners, this specific page 'Al Khuwa Al Nadifa' is known for sharing news pieces and adding polarising (sometimes irrelevant) content to the news piece (Ps: the martyr is an activist who participated in the protests). Finally, the other piece of news that was shared in different posts by different posters is about the parliamentary session to nominate the new President and the subsequent events that happened during.

In Kurdish, these four peaks were:

#### November 28 2021, with 952 posts

7 out of the 10 posts that gathered the most interactions on that day, included the topic 'Ethnic identity Kurdish', and 6 included the topic 'terrorist organization' and 'violence' and 'hatespeech'. A closer look at the posts shows that most posts are about ISIS in Kurdistan and tensions between the Peshmerga<sup>2</sup> and ISIS. These 4 topics were visibly the most recurring together.

#### February 24 2022 with 1,255 posts

The post that gathered the most interactions on this day was from Xendan and was about Russia's war on Ukraine. All 9 remaining posts (from the top 10 based on number of interactions) are about Russia's war on Ukraine as well, including photos of Ukrainians preparing for conflict.

#### February 27 2022, with 1,116 posts

All top 10 posts based on level of interaction on this day are also about Russia's war on Ukraine with a specific bias articulated clearly in captions on news accounts in the Kurdish scene.

#### March 13, 2022 with 1,048 posts

All top 10 posts based on level of interaction on this day are also about Russia's war on Ukraine and were mostly shared by Xendan, Rudaw and Lahur Sheikh Jangi Talabany. Important to note here that Xendan and Rudaw were the two accounts to discuss & engage the most on the war on Ukraine.

# **ARCHETYPES OF POLARIZATION**

Three archetypes of polarization are evident from this analysis.

First, the analysis of the Arabic Facebook sphere shows signals of interest polarization with regards to identity. Specifically, the use of posting tactics that deliberately frame issues to attack an outgroup (Kurds) or exalt an ingroup (Shiites). This is especially clear on pages such as Al Khuwa Al Nadifa, which chose to inform the public and pose questions that clearly transform issues from specific to general, and enhance morality or moral authority. Programming that identified and directly addressed the most viral pages posting in this manner could be considered.

Second, in the Kurdish Facebook sphere the initial hypothesis of participants that hateful narratives would centre around ethnic identity was proven wrong - conversations about Kurdish identity are largely positive, and there is limited engagement on posts about other ethnic identities (unlike in the Arabic sphere). However, a qualitative analysis of the language used in posts reveals a toxic and polarizing atmosphere against religious minorities (especially Yazidis and Christians), fuelled by misinformation and dehumanization, signaling both attitude and interest polarization. Posts that are related to religious ceremonies (especially during religious festivities) draw discriminatory comments and mockery against religious symbols. In addition to discriminatory comments, Islamic verses and sayings of Prophet Mohammad are posted in the comment sections; these verses are not inherently problematic, but are being instrumentalised to support hateful narratives. The network of Facebook commenters shows that the commenters are normal individuals with their original identities, with no remarkable connection that could be drawn between the accounts. A unifying characteristic of comments is the use of plural "we" and "you" - in reference to superiority. As a result of these insights on attitude and interest polarization, Kurdistan24 is changing the strategy on their own page, by training reporters to share content differently and creating a dedicated group of moderators who intervene in comment threads where misinformation and dehumanising narratives towards religious minorities is shared. Further programming in this line could be supported.

Finally, in both Arabic and Kurdish, political actors were the actor type that received the highest proportion of 'love' emoji reactions (over 50% of all their posts). One likely explanation for this is that most people engaging with these posts are unequivocal supporters: where these posts do not result in confrontation, the comments validate in-group values uncritically. This type of affiliation polarisation may seem harmless, but can also be a signal of stifled debate or a "spiral of silence" that comes from self-censorship. Programming that explores the role of political actors in creating a culture of "affiliation polarisation" could be supported.



Produced by Build Up for the DMAPS program

Lebanon Research Report



This research report is based on analysis of 131 Facebook pages and groups and 90 Youtube channels for Lebanon, Palestine, and Syria. These pages, groups and channels were chosen by the four Digital MAPS program partners in the three cluster countries as the places on social media where they believed most polarising conversations around gender stereotypes are taking place. Posts and comments collected from these accounts between November 2021 and March 2022 were labelled by topic (based on the text) and actor (based on who posted). The list of 17 topics and 10 actor types developed by the program partners is available <u>here</u> along with the list of data sources.

This report outlines insights about actors and topics based on content from these 131 facebook pages and groups and 90 Youtube channels. Over the period of November 2021 to March 2022, we reviewed a total of 14,547 relevant<sup>1</sup> Facebook posts and 3,379 Youtube posts. The report also looks at 165,347 scraped comments from Facebook (from a total of approximately 8 million comments). Where this report is based on a comprehensive analysis based on the selected data, it is important to remember that these insights do not reflect the entire social media landscape in Lebanon, and are presented as part of a cluster dashboard that depicts data from the three countries, in addition to being specific to the research statement that the cluster chose at the beginning of their analysis process.

#### **ACTOR MAPPING**

Most of the actors talking on Facebook and YouTube are based in Lebanon with a reach spanning across the MENA region. The top posters are local news sites and media personalities that tend to be prolific and consistent content producers, even when looking at the data across the three cluster countries. When isolating for Lebanon actors only, some of the local media sites such as MTV Lebanon or media personalities like Malek Maktabi remain on the list, and are joined by alternative news sites such as Megaphone. This alternative media outlet, which gained more popularity and online traction since the 2019 protests in Beirut, relies solely on social media for its reach and exposure.

When filtering for topics like women, harassment, and gender based violence, the Lebanese NGO Kafa is at the top of the YouTube engagement list with 11 videos, compared to other major media outlets like Aljadeed with 7 videos uploaded throughout the research period. Kafa's position here shows as an actor working directly on women-related issues and GBV, while Aljadeed uses YouTube to publish content on this polarizing topic as news stories or special reports to be watched after airing. The types of actors posting most frequently are "cultural" actors, a label that encompasses news sites, entertainment pages, and public groups like "The Lady", a community and lifestyle page that targets women and encourages them to post about everyday experiences.

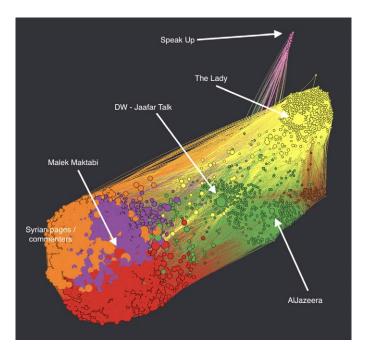
Even without filtering for only Lebanese actors, the pages Mada (an alternative youth-led news and activism network based in Beirut) and Sharika wa Laken (a feminist page publishing news and opinion articles by and for women) both still make the top five pages by number of posts out of the full data set, with 549 and 322 posts respectively. They are not, however, in a similar ranking when it comes to comments. Comment numbers are much higher in general, with Malek Maktabi (over 600,000 comments), The Lady (more than one million comments), and Jaafar Talk (almost two million comments) in the top five. The Lebanese talk show host Malek Maktabi showing in the top five is consistent with the "scoop" and teaser culture that his TV show adopts to gain traction and start conversations about polarizing topics online as a parallel engagement strategy to increase TV viewership.

<sup>1</sup> Posts of relevance in this report refer to posts that are labelled under one or more topics.

Alternative media outlets such as Megaphone and Thawra Memes are frequent posters, but do not receive equal attention in their comment sections. In fact, they are isolated from the dataset and have a negligible amount of comments, such that they do not show up in the visualization. This shows that their large following is passive in essence and does not comment on their posts in general, as they are used as a news source with few alternatives in Lebanon. This also shows a notable polarization between followers of these two pages and the rest of the sources in the dataset, even other Lebanese media outlets.

There is little overlap between the top actors on Facebook and those on YouTube. It is noteworthy that when looking across the three cluster countries, the top YouTube posters are all TV channels or personalities, representing either official channels like AlJadeed, or personalities like Malek Maktabi. Interestingly, this does not change significantly when filtering for only Lebanon-based actors, showing that Lebanese news sites are generally more prolific on YouTube than channels from Syria and/or Palestine, and engage viewers from all three countries whereas the opposite may not be true based on this dataset.

As shown in the captured visualization below, some Facebook actors are central to online conversations and engagement, and lead to certain echo chambers forming around multiple poles. On one side there is The Lady, whose followers are engaging with one another, and to a lesser extent, with followers of other pages in the dataset. The Lady is catered for women only and publishes general conversation starters under the description, "The Lady is a supporting page for women worldwide and specially in the Arab world, we give advice about how to stay positive, be strong and of course fashion and make up." The commenters and followers of this



#### Network of Facebook commenters

page are farthest away from Syrian pages such as Sikkar Wasat and Sawt Souriyat. This is striking only when juxtaposing with the topics emphasized in this dataset: we cannot assume that pages with similar content in a similar geographic area will engage the same users, and the visualization is enough to show us that Syrian pages are siloed from Lebanese pages in the cluster's joint network public sphere.

Additionally, the page Speak Up, which is also targeting women, has some linkages to The Lady, but we can clearly see that the former's content, which is more oriented towards expressing hardship, "heavier" or more triggering sharing of stories, and documenting of everyday violations women may face, is not appealing to the same audience as The Lady, where conversations are more centred on light coverage of current events, fashion, and cooking. The middle of the figure shows us that prolific pages like Jaafar Talk, and media pages like Al Jazeera are wide-reaching in their audience and content. Closer to the edge, we see Malek Maktabi with a far-reaching commenter network. Although it may seem that this page is a bridge builder across multiple echo chambers of commenters, a closer look at the comments and type of engagement on the page shows that the norms of communication on the page are combative, and the topics discussed are posed in a way that invites

engagement but encourages exchanges regardless of the direction of narrative or of language that may be harmful to women or marginalized groups. Some sample comments from Malek Maktabi's Facebook page below.

There is a clear separation between Lebanese media outlets and content producers and those from Palestine. This also presents a lesson learned in the clustering of countries for the dashboard. This difference is much more noticeable on YouTube than it is on Facebook. As mentioned elsewhere, this is also due to the less-engaging nature of YouTube as a search-based platform. The alternative media outlets whom we saw as very siloed in Facebook conversations not only show up on the YouTube graph, but are also closer to mainstream Lebanese and Arab media outlets, and to media personalities with TV shows like Malek Maktabi. It is difficult to discern whether this is due to the nature of YouTube as a search-based platform, or if this is telling of the type of social media users who would look to YouTube for mainstream and alternative news coverage at once.

The right side of the figure below shows more celebrity, gossip, and fashion oriented pages targeted towards women, while the left (green) shows mostly Lebanese media outlets, with pages like Malek Maktabi in the middle of the two. This is similar to Facebook in that topics targeted stereotypically towards women are relatively siloed, far away from harder news topics and sources, with TV shows or actors who aim for click bait and prolific engagement ahead of air time "bridging" the two.

"!أريد الطلاق لأن زوجي لا يلبي كل احتياجاتي"

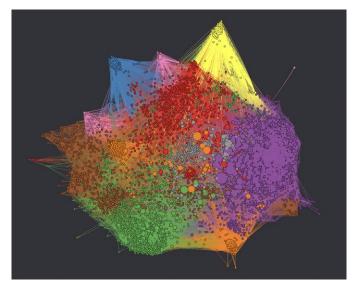
"لست جدّته، أنا زوجته ويرغبني كل يوم"

"أجرت عدداً ضخماً من عمليات التجميل: "أنا إمرأة يرغبها كل الرجال، وهكذا أصبح شكل أنفي

أسمى معاني الأمومة: أم تهز العالم بقصّتها مع إبنها

!أعز من أختى: نتقاسم كل شيء حتى الزوج

!أم تركع أمام الملايين: "أدفع لك عمري لكي أرى أطفالي للمرة الأخيرة"...وتنهار



#### Network of YouTube commenters

On both Facebook and YouTube, we see that there are consistent commenters who are pages and not individuals. A closer look at the dashboard shows a striking difference in the number of posts versus number of comments on the top channels. While videos from the same source could be in the hundreds, such as AlJadeed with 368 videos, the top commenter on YouTube, Lamis Chaaban, shows only 15 comments. It is therefore more useful to conclude that YouTube encourages more passive sharing of news and updates than Facebook, making comment content less relevant to understanding the polarization of the former.

### **TOPIC ANALYSIS**

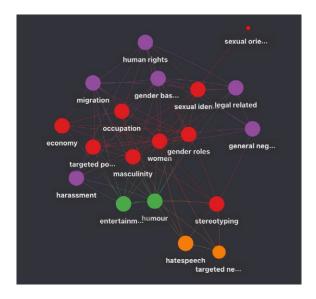
On the cluster dashboard, it is clear that the pages and channels selected by partner representatives were relevant to the topic outlined in the research/problem statement, since the overwhelming majority of relevant posts discussed women, followed by masculinity and gender roles. Other topics appeared at a closer look, but were harder to come across through direct search results, such as stereotypes. Although much of the raw data of the above tagged topics showed an association with stereotypes against women or encouragement of traditional gender roles, stereotypes as a topic did not quantitatively show as one of the most discussed topics online. The quantity of data for only Lebanon actors is telling but should be studied keeping in mind that only pages that were very clearly Lebanese were added to this label. This does not mean that other actors do not discuss issues central to Lebanon, or that other pages are not dominating the Lebanese network public sphere.

While the range of topics tagged did not vary significantly when filtering by actor, the dashboard did show more engagement from NGOs about legal issues, sharing stories about harassment, violations of women's human rights, and success stories about women overcoming daily hardships at work or in public spaces. This indicates a higher likelihood of civil society actors to share such incidents as posts, which is sound within this methodology of only collecting data from public pages.

On Facebook, sexual identity and gender based violence show as top topics for post content, but do not show prominently when looking at comments. On pages that invite engagement on these topics, these data sources show us that actors can safely propose conversations or pose questions on such sensitive issues, but do not receive engagement in the comments as easily. This may be a symptom of an unsafe network public sphere, or a general practice online of people, and especially women, being more cautious about sharing such information in comment sections of public pages.

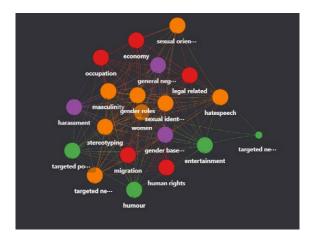
Topics are similar across both platforms, with women, gender, masculinity, and stereotypes at the center of discussions. Interestingly, sexual orientation appears as an outlier topic on YouTube, more so than it appears on Facebook. At a closer look, this is not indicative of silence on the topic of homosexuality or sexual orientation on YouTube. It is rather a sign of less conversation and overall engagement in comment sections on the platform, and a less apparent labelling of YouTube posts relevant to the topic as such.

Also on YouTube, we can see that the topic of stereotyping is a bridge between hatespeech and women, showing that hatespeech on YouTube is likely showing as stereotyping, and that stereotyping is the combative method through which hatespeech is being communicated. Moreover, the topics of migration, gender-based violence, and harassment are linked/spoken about together. A more qualitative look at the data shows that this could be coming from posts from alternative and mainstream media outlets speaking about migrant workers, internal displacement under Israeli occupation, and harassment in the workplace, especially of non-Lebanese workers as part of several online campaigns by actors such as Megaphone, Sharika wa Laken, and Kafa.



#### Network of Youtube topics

The network of Facebook topics shows a clear association between masculinity and "gender identity" or "role of women. This connection is further confirmed in our analysis of wordclouds: the word "woman" or العاسن / أور (لرجر، خيش، جوز), except for the masculinity topic. This gave room for deeper qualitative analysis which showed that men are intervening when women-related topics are discussed (i.e: economic independence, custodial rights, marriage and divorce, gender-based violence), whereas women tend to be sidelined when intervening on other more general topics, some of which are viewed as male-centric in the online space, like political analysis of news headlines or sports.



Network of Facebook topics

The discussion of gender roles and stereotypes generally has a negative sentiment, where "love" and "sad" reactions are the most prominent, which indicates empathy when reacting to such topics, however, "Haha" reactions are also noticeable, indicating sarcasm or a negative sentiment towards the sensitive or harmful information shared in the original post. Based on sentiment analysis, women are more empathetic on Facebook. They use the love reaction on the chosen topics significantly more than men. However, men tend to use the angry and laughing reactions more than women. We can deduce that gender or women-related posts are frequently ridiculed or generate strong negative reactions from men. This is usually conveyed with hate speech, aggressive language, and religious references.

Finally, there are more positive sentiments on Youtube than there are on Facebook, from both men and women gender defined actors. This is highly dependent on the intentionality of YouTube as a platform; the scheme on Facebook is more randomized than YouTube, whereas YouTube is a search-based platform.

#### TIMELINE ANALYSIS

There are few specific overlaps between peaks in the timeline of the data and offline events that affected Lebanon on the whole.

One peak was viewed in mid-January 2022 with concentrated comment traffic on Jaafar Talk, The Lady, and other news/entertainment sites especially on Facebook. This applied to Lebanon in specific and to the entire cluster, as it coincided with the release of the movie as7ab wala a3az or Perfect Strangers on Netflix. The Arabic production of the film included a cast of famous actors from Egypt and the Levant, and came under scrutiny for showing the Egyptian actress Mona Zaky in scenes that were viewed by many as inappropriate, overtly sexual, and against "family values".<sup>2</sup> This was a very polarizing topic on mainstream media and gained immense traction online during this period, especially on pages that invite engagement on "controversial" topics like women's sexual freedom and homosexuality.

February showed no significant spikes in engagement, only signs of added advertising and love-related content around Valentine's day which is consistent with a cluster of topics for the dashboard, and with the approach of many women-targeting pages in the dataset.

We can see a major peak on Facebook around international women's day on March 8. This does not coincide with significant offline events, also showing the ceremonial nature of the occasion, and the intentional content production and publishing on this day. A zoom in of the raw data shows content that is stereotypically aligned with women's conventional roles, and this is visible on Facebook not on YouTube.

### **ARCHETYPES OF POLARISATION**

From the analysis above, there is evidence of polarizing attitudes on social media posts discussing the roles women can play in society, where there is clear evidence of a perceptual shift towards stereotypes. This is clear not just in a qualitative analysis of posts, but also in observing the topic network graph, which shows that the issue of the role of women is largely addressed in terms of masculinity and gender identity. Digging into these posts showed that conversations revolved around the appropriate jobs for women based on physical condition or customs and traditions. On the other hand, the role of women is not often discussed in relation to human rights or (even less) economic issues including material needs.

This polarization of attitudes towards women has further resulted in the perpetuation of polarized norms in the network public sphere. Men have a greater tendency to stereotype women, as was clear in qualitative analysis showing frequent comments that suggest emotions are a sign of weakness in women, or sentiment analysis showing negative reactions towards posts indicating that women have financial and educational independence. On posts addressing issues of gender on Facebook, women use mostly love reactions where men use angry and laughing reactions, suggesting that men tend to ridicule this content. As discussed above, wordclouds show a pattern where the word "woman" or  $d_{\alpha}/$ 

In general, participants observed that men either react negatively to content about women (quoting religious ideas or arguments) or approach it with ridicule (stereotyping women by using humor).

The analysis results on polarised attitudes and norms around gender informed the content strategies of the Lebanese team on this programme, who have designed an approach that uses humor and gender conscious content to mainstream gender equality and break gender stereotypes. Further programming to challenge norms and attitudes might address issues of masculinity more directly, including by enlisting public male figures who could call in some of this harmful polarization.



**DMAPS Reports** 

