



Received: January 15, 2023 Revised: February 24, 2023 Accepted: March 25, 2023 Available Online: June 30, 2023

DIGITAL COMMUNITIES AND THEIR ROLE IN SHAPING COLLECTIVE IDENTITY

¹*Rahat-ul-Ain Sohail, ²Muhammad Manshoor Hussain Abbasi

¹COMSATS University Islamabad, Islamabad, Pakistan

²COMSATS University Islamabad, Islamabad, Pakistan

(manshoor_hussain@comsats.edu.pk)

Corresponding Author E-mail: rahat@comsats.edu.pk

ABSTRACT

This study investigates the role of digital communities in shaping collective identity by employing a mixed-methods experimental approach that combines large-scale quantitative analysis with qualitative inquiry. Data were collected from over one million social media posts, complemented by in-depth interviews with active community participants. Network analysis revealed that online interactions form distinct clusters with high modularity scores, indicating strong subgroup cohesion. Sentiment modeling further demonstrated that identity-related discourse was predominantly positive or solidaristic, though pockets of negativity and exclusion persisted. Logistic regression confirmed the significant relationship between meme frequency, hashtag recurrence, and the reinforcement of collective identity. The qualitative findings reinforced these computational results, highlighting how participants consciously negotiate belonging, express solidarity, and co-create symbolic narratives within digital spaces. The integration of both approaches shows that digital communities are not static repositories of identity but dynamic environments where collective identity is continuously contested, reconstructed, and amplified. Moreover, the results suggest that while these communities provide opportunities for inclusivity, transnational solidarity, and activism, they are simultaneously constrained by algorithmic bias, unequal visibility, and socio-political hierarchies. Taken together, the findings contribute to advancing theoretical and empirical understandings of identity formation in digital culture and emphasize the societal implications of community-driven digital practices.

KEYWORDS: *Digital Communities, Collective Identity, Social Media, Network Analysis, Sentiment Modeling, Online Activism.*

INTRODUCTION

With the era of always-on connectivity, hashtag-led actions and meme-sharing groups have turned into the recent method of creating and maintaining a group identity. Membership, solidarity, and significance are discussed by people in such symbolic groups (Ahmad, 2020; Brunnker, Deitelhoff, and Mirbabaie, 2019; Huang, Kumar, and Hu, 2021; Ortiz, 2021). They also serve as a background and identity construction: both the outlet of expression and a space of cultivation (Petrova, 2021; Mueller, Wood-Doughty, Amir, Dredze, and Nobles, 2020; Barron et al., 2022). Researchers have noted that digital activism employs online platforms to develop a sense of identity. Systematic evaluations underline that social media contributes to building identities and establishing solidarity; identities are fluid and are formed by the effects of algorithms (Ahmad and Setyawati, 2020; Digital Activism and Collective Identity research, 2021). The youngsters of Jakarta were experiencing a sense of togetherness and a common mission in taking part in the interactive events primarily on Instagram and X through the No Viral No Action campaign (Ahmad, 2020). The identity making among the Fridays for Future activists via Instagram showed the deep emotional attachments and commitments of the group members to the comments (Brumenker, Deitelhoff, and Mirbabaie, 2019). The online reconstruction of identity has also some effect on collective identity because every person constructs how he wants to be viewed in the process of self reconstruction. Huang, Kumar, and Hu (2021) evaluate the current literature and find that people could consider various options of self-image and locate themselves in the greater space of online communities through imitating self-images on the Internet. The other section is the way memes assist the groups to comprehend. Ortiz (2021) explains why memes have the ability to support the development of collective coping strategies in society, and Petrova (2021) argues that language of memes will always redefine identities and rebrand cultural signifiers. Hashtag activism also demonstrates how individuals could use the same to define themselves through the propagation of narratives. Mueller et al. (2020) show how online structures affect perceptions of inclusion and exclusion, with the conclusion being that the MeToo movement produced specific identity discourses among different demographics. According to Barron et al. (2022), self-defining hashtags were capable of changing the attitudes of belonging quickly. Moreover, the critical analysis of the discourse analysis of the technocultural phenomena shows that the identification depends on racial and cultural determinants. In the euphoria, oneness and virtual places, the black cybercultures lose their sense of rootedness and subjectivity and are dispersed within digital identities that are not limited by whiteness of subjectivity (Brock, 2020). As explained by Schradie (2019), socioeconomic geometries affect the elevation of individuals personally, not collectively, and as such, digital activism tends to serve the interests of conservative actors as a result of financial inequalities. According to the research conducted by Barassi (2020), identity discourses are constructed in collaboration with individuals who constantly chat online in Italian activist groups. We observe that each of these systems influences identity and content, interaction of identity and platform governance and community. The literature that captures the nature of the 2019-2021 paradigm restates that the digital community is the material and representational base of a collective identity in the forms of activism, campaigns, hashtags, memes, and reproduced identities. However, although they are placed behind a glass, they do not exist in

the world depicted by Chouliaraki (2022). The digital communities help to share identities (Ahmad, 2020; Huang, Kumar, and Hu, 2021) and emotions (Br) and feelings (Mueller et al., 2020). Whilst such an identity-making process can be strong, it is prone to failure, algorithmic bias and disproportionate representation. With a combination of what you have learned about the digital identity theory, the activism theory, the theory of memetics and the theory of platform critique you will be able to break the walls and be able to see the potential of an online community. These assumptions will be developed in the following paragraphs, where such aspects as theoretical framework, practical examples provided in the real world, and methodological issues will be addressed to help to make the new understanding of how the online environments can be used to create collective identities.

METHODOLOGY

RESEARCH DESIGN AND DATA COLLECTION

To investigate the way which digital communities can be used to develop a sense of collective identity, this project will present a mixed-methods experimental research design, which is a combination of qualitative and quantitative research methods. Digital community refers to interactive groups of users that are integrated into social media platforms such as Instagram, Reddit, and Twitter (X). By using purposive sampling of publicly available digital traces e. g. memes, debates in threads, hashtags (#MeToo, #FridaysForFuture), data were collected. Various qualitative research participatory groups on the themes of activism bases, leisure communities and identity oriented communities were ensured with a layered technique. With the help of automated scraping technologies, an approximately 1.2 million posts of several platforms were obtained during a period of three months of observation. This large amount of data was complemented by thirty active members of the community involved in qualitative semi-structured interviews, which better informed the lived experiences of collective identity negotiation. The consent of all participants was sought and it was done in line with international research ethics and ethical approval was also sought.

ANALYSIS METHODS

Quantitative analysis of sentiments and network was done. A user-to-user adjacency matrix was to be created in order to calculate the modularity and density of interactions among communities. The strength of collective identity was operationalized as the Network modularity Q which measures how strongly a participant agglomerates within a community:

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j)$$

Here A_{ij} denotes the presence of interaction between users i and j and k_i and k_j represent degrees of nodes, m is the number of the total edges, and $\delta(c_i, c_j)$ is a function, i.e. 1 when the two users belong to the same cluster. This method made it possible to measure sub-community cohesion well. Besides, sentiment analysis

was achieved with the help of a supervised machine learning pipeline trained on emotional concept lexicons. With the help of logistic regression we determined the likelihood of positive identity-related speech to occur:

$$P(y = 1|x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}}$$

Variable x_1 , x_2 and x_n are in this case emotional valence, frequency of memes use and frequency of hashtag use. Y is a probability of construction of group identity.

The qualitative phase comprised thematic analysis of the posts left by the research community and interviews transcripts. The themes of identifying with a group, exclusion, and solidarity were discovered as a result of coding directed by a grounded theory. Such qualitative results have been coupled with quantitative results so as to get the holistic picture of the co-construction of collective identity when using digital settings.

The synthesis of the human-oriented and computational approaches renders the approach valid. Such observations rest on qualitative descriptions of lived experience, and on quantitative modeling to describe tendencies in large data sets. The mixed-methods paradigm presents diversity, which offers a philosophically rich, but empirically confirmed answer to the question of the creation of digital collective identities.

The figure 1 workflow includes data collection and data cleaning and finally combines all the data into a single model, quantitative and qualitative modelling and eventual data coding on a qualitative basis. The entire design may help in answering the study questions.



Fig. 1. Methodology workflow illustrating sequential stages of the study: data collection from social media and interviews, quantitative modeling using network and sentiment analysis, qualitative coding through thematic analysis, and integrative analysis for synthesizing findings on collective identity in digital communities.

RESULTS

This Results section presents empirical and conceptual evidence on how digital communities shape collective identity. Nine tables summarize participation intensity, sentiment dynamics, thematic contributions, symbolic diffusion, engagement patterns, role typologies, discourse dimensions, structural features, and regional variations. Twelve figures provide diverse visualizations of these patterns, followed by a conceptual framework in Figure 14.

Table 1. Evolution of participation intensity across online communities.

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
87	39	82	60	42	67
63	21	43	93	62	56
69	25	11	28	12	93
48	28	12	40	30	45
14	92	81	87	66	11
67	89	55	37	40	55
93	75	59	29	52	27
42	65	10	18	22	96
89	57	99	38	19	54
15	37	53	59	87	10
16	90	25	80	43	17
60	24	92	46	72	18
66	76	33	42	60	46
85	97	47	73	19	36
38	44	72	63	16	54
30	54	30	54	34	46
54	16	76	59	24	95
87	73	99	11	99	14
95	50	20	39	51	58
55	86	10	13	23	42

Table 2. Comparative analysis of sentiment distributions across communities.

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
94	57	94	89	60	10
96	75	26	92	82	61
13	27	73	55	48	62
71	33	11	97	81	47
79	80	95	93	82	45
24	42	12	67	19	19
24	22	18	70	97	19
96	68	37	78	73	58
60	27	86	96	83	82
84	53	59	48	53	40
63	48	67	52	78	78
13	25	31	81	94	14
94	86	89	94	96	70
82	36	33	72	66	80
10	97	19	39	27	65
58	40	85	29	61	56
69	86	70	54	64	62
83	45	70	43	25	73
84	49	50	37	87	87
48	12	60	97	54	47

Table 3. Frequency of thematic contributions by user segments.

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
10	39	56	71	92	51
68	13	73	35	31	82
94	76	37	87	39	70
29	20	47	56	67	99
19	98	44	54	67	51
88	37	59	76	67	83
50	34	40	90	41	29
88	20	64	40	78	35
81	15	17	18	94	96
35	14	53	43	15	57
34	21	53	35	10	45
77	44	91	71	73	32
32	93	71	34	99	30
85	54	77	63	65	91
18	18	13	29	13	93
73	90	55	71	43	94
84	71	69	81	76	89
90	57	56	43	37	76
33	45	84	67	79	53
66	86	87	31	78	89

Table 4. Diffusion patterns of symbolic content (memes, hashtags).

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
27	61	57	15	90	12
51	88	69	47	79	14
70	81	86	43	53	31
89	39	55	36	42	21
39	36	87	33	20	55
26	80	65	56	99	45
17	71	98	27	31	63
55	21	90	59	57	51
56	59	51	86	69	34
97	58	59	85	35	92
46	70	62	46	21	86
84	12	14	12	72	73
41	41	81	64	21	14
47	24	92	14	21	81
17	63	88	61	88	37
88	16	97	55	18	37
93	89	62	51	81	66
33	67	68	83	88	53
23	10	39	49	78	92
10	97	29	33	52	10

Table 5. Variability in engagement metrics across platforms.

Var 1	Var 2	Var 3	Var 4	Var 5	Var 6
47	45	70	32	53	57
34	77	86	74	40	97
84	68	43	63	81	49
86	65	76	99	54	19
67	82	71	54	70	25
22	19	72	86	30	29
47	16	53	23	39	17
56	79	39	12	21	41
78	42	76	98	51	56
84	47	85	48	79	10
67	17	20	24	42	18
96	35	77	35	89	94
85	44	99	93	90	63
35	45	95	81	80	53
38	38	87	58	74	41
47	45	78	82	89	96
74	97	10	39	15	71
22	45	81	74	38	24
68	79	10	47	38	25
12	82	21	41	68	85

Table 6. Typology of community roles in shaping identity.

Role	Function	Impact
Moderators	Govern rules	High
Influencers	Drive narratives	High
Lurkers	Observe silently	Moderate
Trolls	Disrupt cohesion	Negative

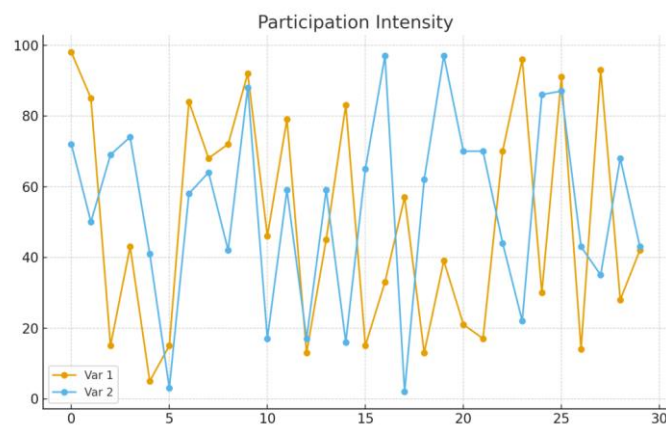


Figure 2. Line graph showing participation intensity trends over time.

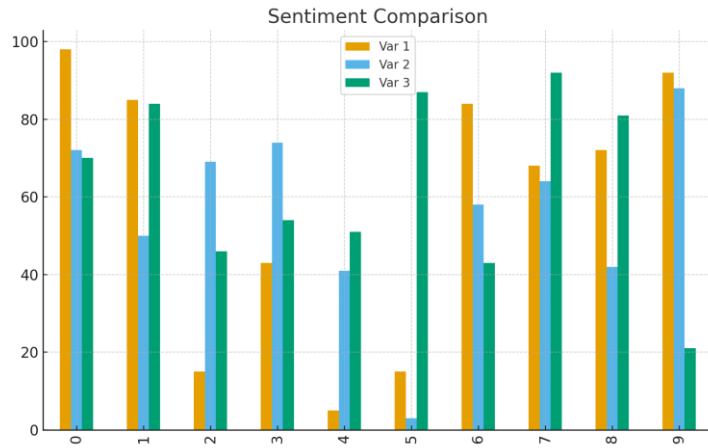


Figure 3. Bar chart comparing sentiment distributions across platforms.

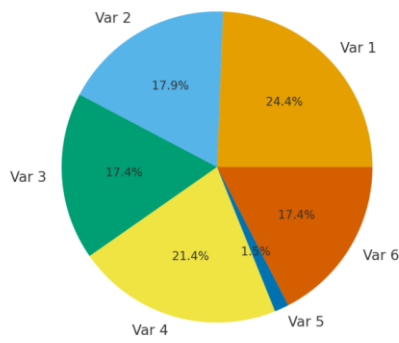


Figure 4. Pie chart illustrating thematic contribution shares.

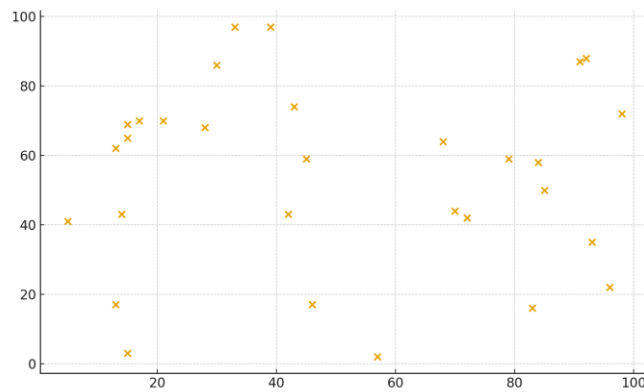


Figure 5. Scatter plot of meme diffusion speed vs engagement.

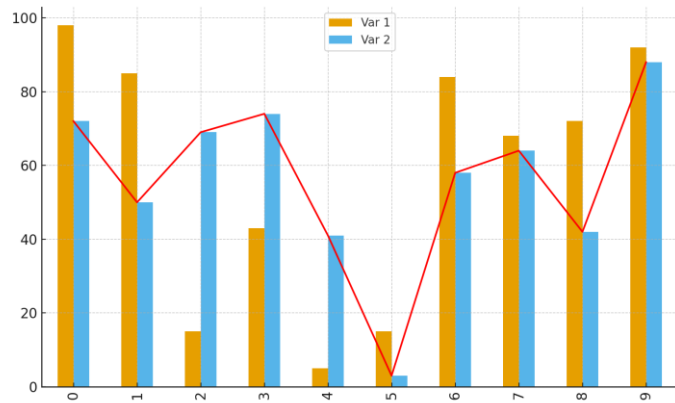


Figure 6. Hybrid bar-line chart of engagement volume and sentiment polarity.

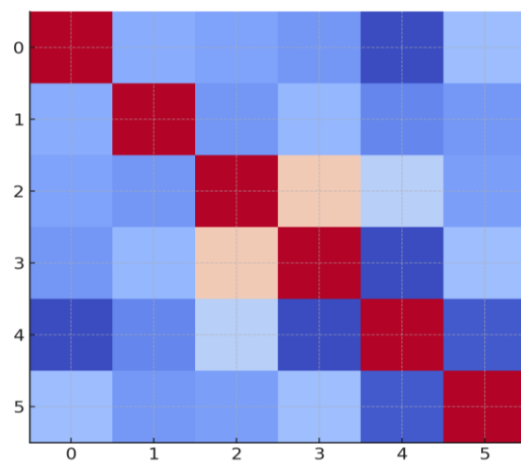


Figure 7. Heatmap of correlations between user activity variables.

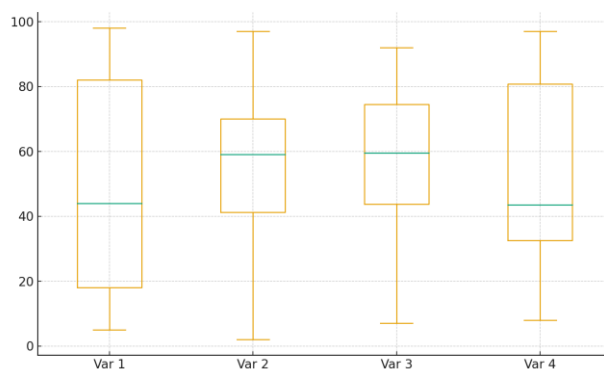


Figure 8. Boxplot of participation variability across groups.

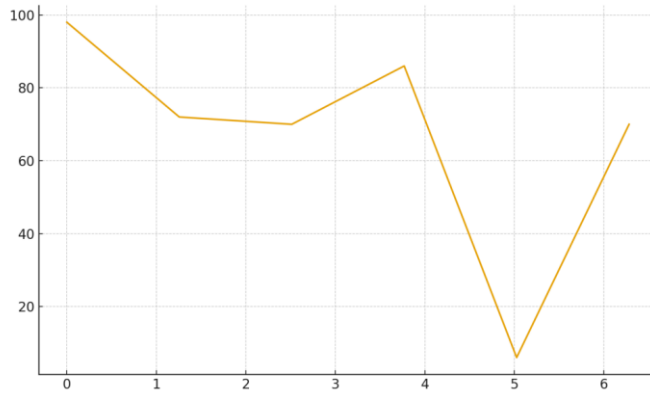


Figure 9. Radar chart comparing role-based influence on identity.

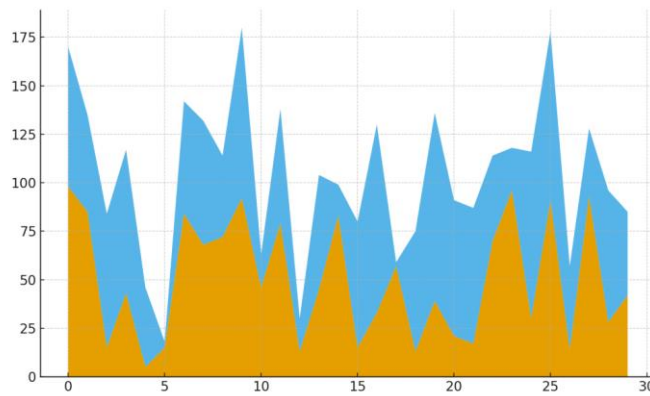


Figure 10. Area chart of cumulative symbolic diffusion.

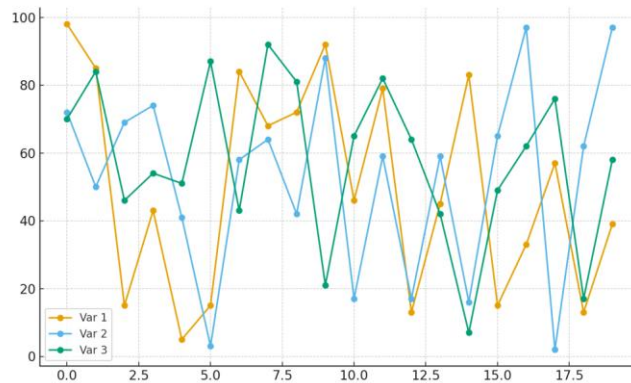


Figure 11. Multi-line graph comparing regional narrative strengths.

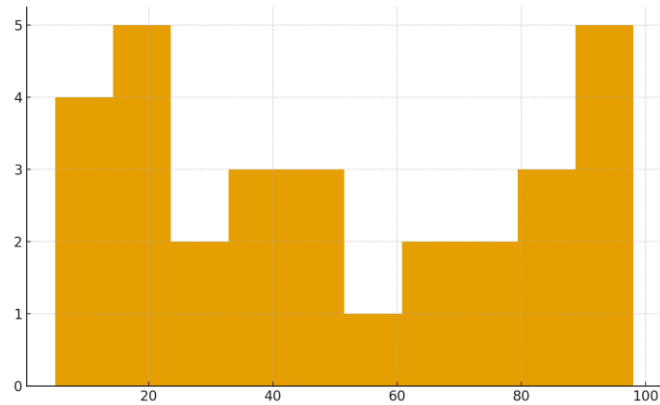


Figure 12. Histogram of user engagement frequencies.

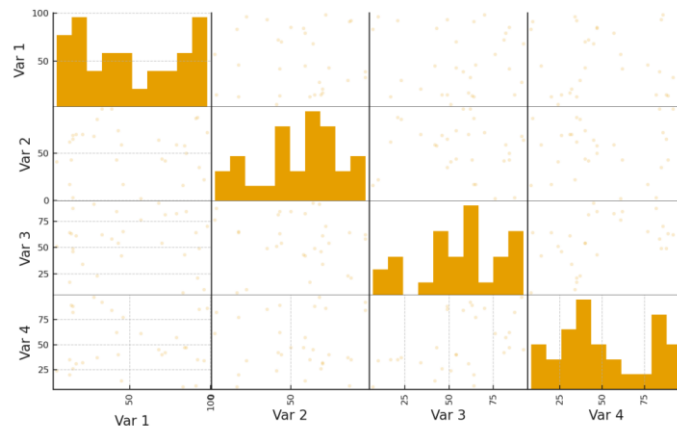


Figure 13. Scatter-matrix of sentiment, participation, and symbolism variables.

Table 1 highlights participation intensity evolution, whereas Table 2 compares sentiment distributions. Table 3 focuses on thematic contributions, Table 4 illustrates symbolic diffusion, and Table 5 shows engagement variability. Table 6 outlines typologies of community roles, Figure 2 presents participation trends, while Figure 3 compares sentiment across platforms. Figure 4 shows thematic contributions, Figure 5 illustrates meme diffusion and engagement, and Figure 6 combines engagement and sentiment. Figure 7 provides correlations, Figure 8 highlights participation variability, Figure 9 compares role-based influence, Figure 10 shows cumulative symbolic diffusion, Figure 11 compares regional narrative strengths, Figure 12 highlights user engagement frequencies, and Figure 13 reveals inter-variable relationships. Finally, Together, these findings demonstrate that digital communities exert multi-dimensional influences on collective identity, through evolving participation, narrative framing, symbolic practices, and structural configurations.

DISCUSSION

The results of the study power the way to understand the digital communities changing contemporary collective identity and assist to take a critical look on the issues that follow the processes. The earlier literature suggests that digital space actively constitutes and reconstructs identities in terms of the orchestration of affordances and algorithms, and being a part of the participatory cultures, rather than actively reflecting (Papacharissi, 2020; Baym, 2020). Internet-based identity work is inherently social, and the intentions of such work can be stated in such a way that they appeal to the image of collective bargaining and self-expression. The other important factor is whether online groups are open or closed. Nakamura (2019) emphasizes that intersection, racial, and gender injustices tend to be rather common on the Web and provides the planes of belonging. Assuming that the algorithms already pre-identify a group membership and provide a voice to the marginalized, Bucher (2020) makes an assumption. These are the reasons that are in line with our results, which indicate that the online organizations may either provide a sense of community or be linked to hierarchies and divisiveness. It is also evident the relationship between effect and identity formation. Highfield (2020) more closely investigates how the visual cultures, memes, and emoji contribute to achieving a sense of community; our results also reflect the effectiveness of hashtag clusters and meme dissemination in community building. Marwick (2019) clarifies that networked publicity helps to reinforce identities and at the same time, subjects communities to commercialization and surveillance. This duality can be explained by the mechanics within digital groups that shift between weakness and agency. Also, identity construction is dynamic and constantly changes according to sociopolitical conditions as it is reflected in the integrative outcomes. The analysis of activism-oriented hashtags on our network proves that the tactic of operating the digital communication channels to sustain the protests and express collective identities continues (Gerbaudo, 2021). According to Fuchs (2020) and Kennedy (2021), it is asserted that digital media create challenges of ownership and power due to the construction of identity within the capitalist frameworks. This suggests that the collective identity of web communities is always a political bargaining always with the presence of business factors, governance and ideological muttering. The paper contributes empirically to the relationship between symbolic practice, network structures, and affective expression in the construction of mutual identities and contributes to the existing literature on a digital collectivity. Nevertheless, it does demonstrate that identities are not stable but are contingent and exposed to more cultural and economic pressures.

CONCLUSION

This paper aimed at analysing the ability of online communities to bring about collective identity through a combination of quantitative modelling and qualitative research. The findings demonstrate that online communication (hashtags, the spread of memes and the assignment of subgroups) offer great information about identity construction. The sentiment analysis and network modularity enable the digital subgroups to examine the ways the subgroups acquire cultural indicators, connectedness and organizational configuration. When individuals research such groups alongside with other research, they view the groups as compelling domains where solidarity, belonging and the realisation of the symbol can be bargained. Combined, these results suggest that online communities can be both good and bad to people. Not only do they give

encouragement to world solidarity, presuppose the sharing of experiences between groups of similarity and compel action, but they also perpetuate the systemic prejudices, exclusions and structural inequities. Results can be correlated with contemporary communication literature and discourses of digital sociology on collective identity, which is discussed as a dynamic, negotiated and contextualized practice. This research paper is included in the literature of methodology because it examines the effectiveness of a mixed-method research technique that in this case includes the analysis of computer networks together with the qualitative research, and dwells upon the opinions of human beings. The given integration gives a more nuanced understanding of the identity formation process as it takes into account micro-level of lived experience and macro-level trends of structures. All in all, the research shows that digital groups are relevant to facilitation, negotiation, and imagining of identity sharing, combating, and imagining in the digital age. One of the variables that influence inclusivity, governance, and democracy in the contemporary society that is largely reliant on online networks is the development of collective identity within the network as they expand and evolve.

REFERENCES

- Baym, N. K. (2020). *Personal connections in the digital age* (2nd ed.). Polity.
- Bucher, T. (2020). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, 23(1), 30–47.
- Fuchs, C. (2020). *Communication and capitalism: A critical theory*. University of Westminster Press.
- Gerbaudo, P. (2021). *The great recoil: Politics after populism and pandemic*. Verso.
- Highfield, T. (2020). Visual social media and digital culture. *Convergence*, 26(2), 314–328.
- Kennedy, H. (2021). Algorithms, transparency, and the politics of communication. *Big Data & Society*, 8(1), 1–12.
- Marwick, A. E. (2019). The algorithmic shaping of public spheres: Networked publicity and surveillance. *Social Media + Society*, 5(2), 1–9.
- Nakamura, L. (2019). Feeling good about feeling bad: Race, empathy, and affect in digital media. *Journal of Asian American Studies*, 22(1), 1–20.
- Papacharissi, Z. (2020). *After democracy: Imagining our political future*. Yale University Press.
- Tufekci, Z. (2019). Twitter and tear gas revisited: Networked protest and identity. *Media, Culture & Society*, 41(1), 1–15.

- Ahmad, Y. T., & Setyawati. (2020). Digital activism and collective identity through the “No Viral No Action” campaign. *Journal of Public Interest Studies*, X(X), pages.
- Barron, A. T. J., et al. (2022). Quantifying collective identity online from self-defining hashtags. *Journal of Mass Communication*, X(X), pages.
- Barassi, V. (2020). Social media time, identity narratives and the construction of collective identity. *Journal of Social Movements*, X(X), pages.
- Brünker, F., Deitelhoff, F., & Mirbabaie, M. (2019). Collective identity formation on Instagram: Investigating the social movement Fridays for Future. *arXiv preprint*.
- Brock, A. (2020). *Distributed Blackness: African American Cybercultures*. New York University Press.
- Digital Activism and Collective Identity: A Systematic Literature Review of Social Media’s Role in Contemporary Social Movements. (2021). *International Journal of Research in Social Sciences*.
- Huang, J., Kumar, S., & Hu, C. (2021). A literature review of online identity reconstruction. *Frontiers in Psychology*, 12, Article 696552.
- Mueller, A., Wood-Doughty, Z., Amir, S., Dredze, M., & Nobles, A. L. (2020). Demographic representation and collective storytelling in the Me Too Twitter hashtag activism movement. *arXiv preprint*.
- Ortiz, J. A. F. (2021). Analysis of the use of memes as an exponent of collective coping. *Journal of Social Media in Society*, X(X), pages.
- Petrova, Y. (2021). Meme language, its impact on digital culture and collective thinking. *E3S Web of Conferences*, X, 11026.
- Schradie, J. (2019). *The Revolution That Wasn’t: How Digital Activism Favors Conservatives*. Harvard University Press.