The Emergence of Collective Identities Online: Detecting Patterns in Social Movement Formation on Twitter

Keywords: collective identity, digital trace data, socio-semantic network analysis, social movements, social media

Extended Abstract

It is undeniable that online actions of social movements such as #MeToo affect the offline world by, e.g., impacting public opinion (Szekeres et al., 2020). Therefore, it is vital to further the understanding of the formation and development of movements online. Yet, demarcating any online group from its environment and monitoring its dynamic evolution has been a long-lasting problem in research, particularly in social network analysis (Chouchani and Abed, 2020). Several studies have addressed this issue through a methods-driven lens, e.g., relying on network characteristics like community detection algorithms (for an overview, see, e.g., Dakiche et al., 2019). However, to detect and track their structural development and substantially attribute meaning to those groups, we need to combine robust theoretical frameworks with rigorous empirical methods. By doing so, we can fully exploit the rich insights provided by "digital trace data," that is, "records of activity (trace data) undertaken through an online information system (thus, digital)" (Howison et al., 2011: 769).

To account for that issue, we will demonstrate that analyzing group formation through a collective identity perspective is a novel approach to differentiate between groups and track their development on a structural and content-related level. By combining Harrison White's (2008) relational understanding of collective identity with theories from the social movement research (e.g., Melucci, 1996), we propose a novel and generalizable framework for empirically analyzing the formation and development of a group's collective identity online (see Figure 1). We will translate this framework into a socio-semantic network analysis by acknowledging the duality of the social and cultural order of collective identity (White, 2008).

As a second step, it is necessary to test whether the process of collective identity formation follows specific patterns that can be generalized across online groups to systematically track their dynamic evolution based on their collective identity development. Therefore, we apply the framework mentioned above to the two social movements, "Querdenken" (i.e., the most successful group to mobilize against the Covid-19 measures in Germany) and "Fridays for Future" and analyze their collective identity formation over time. By doing so, we test whether (1) the movements develop a collective identity that demarcates them from their environment and (2) if this process can be subdivided into similar overarching stages across movements. Concretely, we answer the following two research questions:

RQ1: Do social movements develop a collective identity that demarcates them as distinct actors online?

RQ2: Can the formation process of social movements' collective identity online be separated into unifying stages across movements?

To analyze collective identity formation online, we rely on the social media platform Twitter, which is, amongst others, highly important for social movement mobilization (e.g., Jungherr, 2015) and provides both interactional and semantic traces that are necessary for our analysis. For each movement, we perform the following five data collection steps (see Figure 2). First,

we collect data related to a general discourse the movement will likely be part of (for Querdenken, e.g., the Covid-19 discourse). As a second and third step, we reduce the dataset to tweets that contain hashtags and have been retweeted since both pieces of information are necessary for our socio-semantic network creation. Fourth, we detect social movement supporters within the data by starting with self-declared movement adherents and expanding this base by manually coding their followee lists regarding movement-related users. Fifth, we divide the remaining dataset into different time intervals corresponding to important events impacting the social movement. Doing so allows us to analyze collective identity formation at different stages dynamically.

For each of those timeframes, we perform a socio-semantic network analysis (for a more detailed overview, see Figure 3) by first combining socio and semantic information (users and hashtags) and, second, analyzing them individually (user networks and hashtag networks). Once the networks are created, we apply the community detection algorithm Louvain to detect communities on a structural basis. Those communities are then analyzed and labeled according to their shared content and the normative evaluation thereof. Both aspects are manually coded based on hashtags and tweets selected via k-core decomposition.

When comparing the structural and semantic development of the communities across movements and over time, our preliminary findings show that a movement's collective identity is characterized by fluid actor constellations and solidifying and increasingly charged meaning structures that demarcate a movement as a distinct social actor (RQ1). We are currently analyzing whether these developments can be separated into reoccurring patterns (RQ2). Since the analyses have shown that actor constellations are volatile across movements and do now allow for detecting reappearing patterns, we focus on shared topics and their normative evaluation. More specifically, we analyze whether emerging topics can be aggregated to broader themes across movements while intensifying their normative evaluation.

Overall, this paper furthers the understanding of dynamic collective identity processes that are crucial for social movement formation online. Thus, our findings add to the growing body of online group detection and tracking by combining a theoretical framework and a thorough empirical analysis.

References

Chouchani N and Abed M (2020) Online social network analysis: detection of communities of interest. *Journal of Intelligent Information Systems* 54(1): 5–21. DOI: 10.1007/s10844-018-0522-7.

Dakiche N, Benbouzid-Si Tayeb F, Slimani Y, et al. (2019) Tracking community evolution in social networks: A survey. *Information Processing & Management* 56(3): 1084–1102. DOI: 10.1016/j.ipm.2018.03.005.

Howison J, Wiggins A and Crowston K (2011) Validity issues in the use of social network analysis with digital trace data. *Journal of the Association for Information Systems* 12(12). Association for Information Systems: 767–797. DOI: 10.17705/1jais.00282.

Jungherr A (2015) *Analyzing Political Communication with Digital Trace Data. The Role of Twitter Messages in Social Science Research*. Springer. DOI: 10.1007/978-3-319-20319-5.

Melucci A (1996) *Challenging Codes: Collective Action in the Information Age*. Cambridge University Press. DOI: 10.1017/cbo9780511520891.

Szekeres H, Shuman E and Saguy T (2020) Views of sexual assault following #MeToo: The role of gender and individual differences. *Personality and Individual Differences* 166: 110203. DOI: 10.1016/j.paid.2020.110203.

White HC (2008) *Identity and Control: How Social Formations Emerge*. Princeton University Press.

Figures

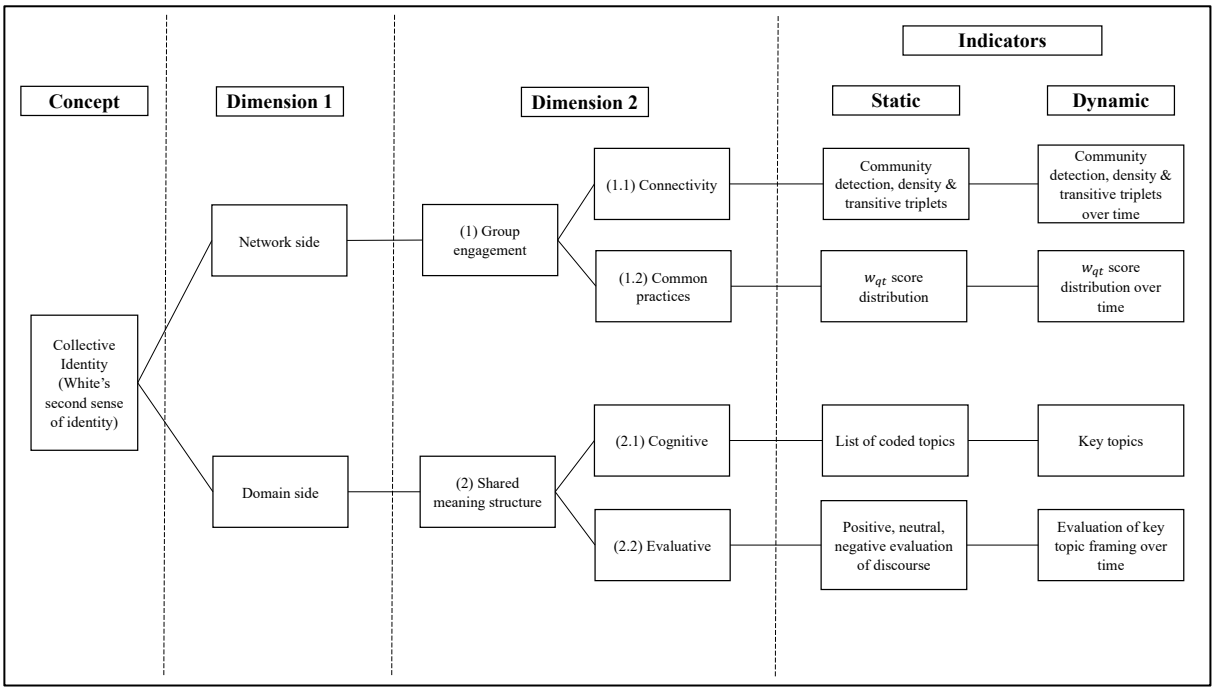

Figure 1. Operationalization of relational manifestation of collective identity online, with empirical indicators.

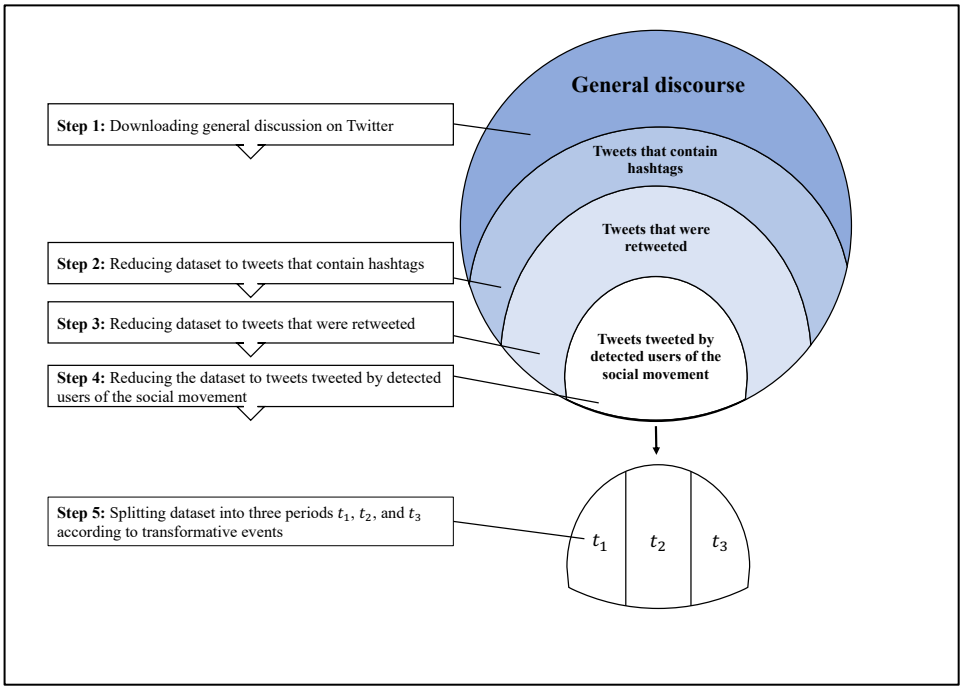

Figure 2. Data collection steps for individual social movements on Twitter.

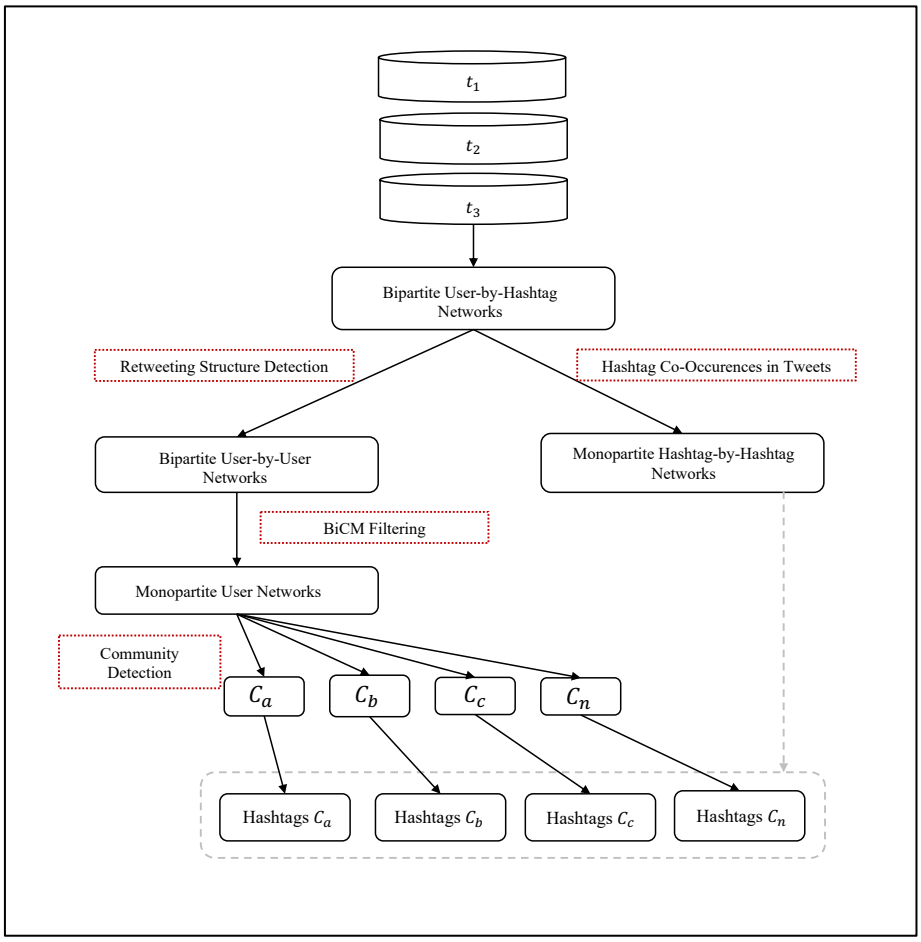

Figure 3. Socio-semantic network creation and community detection.