Image networks and practice analysis of larger data corpora. An approach to cluster and recontextualize visual practice in social media

Wolfgang Reißmann*, FU Berlin, Institute for Media and Communication Studies, Journalism Studies, Germany
Miriam Siemon, FU Berlin, Institute for Media and Communication Studies, Journalism Studies, Germany
Moe Kinoshita, FU Berlin, Institute for Media and Communication Studies, Journalism Studies, Germany
*Corresponding author: wolfgang.reissmann@fu-berlin.de

Abstract
This paper reports a methodological exploration combining image network analysis and standardized practice analysis on social media data. Through applying the open source software Memespector to access the Clarifai API, the potential of an easy-at-hand image tagging tool as an instrument to manage larger data corpora is explored. Using the example of the German-speaking Twitter hashtag #systemrelevant, we relate image clusters to the results of standardized practice analysis of posts that contain images. The proposed method is intended for research that attempts to carve out the co-constituting of public discourse in social media by different groups of actors. The approach systematically differentiates the contributions of societal groups such as journalism, civil society, or private individuals, and the embedding of their tweets in selected anchoring practices and further modalities of participation. Altogether, the multistep analytical process offers a possible approach to process larger image corpora, while maintaining a sensitivity for the practice-theoretical demand of (re)contextualizing image use.

Keywords
image network analysis, multi-methods, large data corpora, visual practice, image clusters, image types

1 Introduction
This paper demonstrates the potentials and limitations of a method which combines automated image network analysis and a standardized coding of the communicative context in which the images are embedded. It is suggested as a method for analyzing larger data corpora which represent the “natural” unfolding of public discourse constituted by heterogenous groups of actors, e.g., based on hashtags or significant keywords. In a first step, we situate our approach in the current landscape of visual communication and social media research. Afterwards, we present our method and its exploration using the example of #systemrelevant, a German Twitter hashtag in the context of the COVID-19 pandemic. To conclude, we discuss benefits, difficulties and limitations we found.

#Systemrelevant contains intense public discussion about which professions are indispensable during crises, how they are socially valued, and by whom they are executed. The hashtag gained momentum in mid-March 2020, when first substantial policies regarding social distancing came into effect. The hashtag mainly thematizes working conditions of health care workers in hospitals and residential care homes. The data are part of a bigger research project which reconstructs the performative making of public discourse on gender-related issues in social media through heterogenous groups of actors. #Systemrelevant was chosen as an example for public negotiation of the societal value of professional care work.

This paper reflects an additional methodological exploration toward a practice-oriented analysis of large(r) image data corpora. We do not put emphasis on gendered visuali-
2 Visuality in social media and practice-based research

A growing body of research is dedicated to studying visuality in social media. The focus is broadly divided into two kinds: images as representations and images as practices.

Research that focuses on images as representations is predominantly media-centered, analyzing the visual and textual content of images. This branch of research either applies established forms of content analysis (e.g., Fahmy & Ibrahim, 2019, for an integrative framing analysis of memes), follows the traditions of iconography and iconology (e.g., Wetzstein, 2017, for a documentary image analysis), or uses combinations of quantitative and qualitative image analysis (e.g., Liebhart & Bernhardt, 2017, for a prolongation of image type analysis sensu Grittmann, 2007).

A common characteristic of approaches studying visuality in social media is “to abandon any idea of analyzing individual images” (Hand, 2016, p. 216). Social media include various visual genres (Highfield & Leaver, 2016): from photos of text quotes and screenshots, to vernacular photography, to press photos, cartoons, memes, art, and more. Therefore, ways of typologizing images or characterizing certain genres gained importance. In the mid-2000s, classifying the spectrum of “profile images” (Astheimer, Neumann-Braun, & Schmidt, 2011; Meier, 2009) of the “new” social network sites using qualitative approaches was pioneering the field. A few years later, image genres such as “memes” (Miltner, 2018) or “selfies” (Frosh, 2015) formed own research fields.

Research interested in image or visual practice adds sensitivity for practices of image production and/or for images “as part of wider embodied practices” (Lehmuskallio & Gómez Cruz, 2016, p. 6). This contextual view scrutinizes images and other forms of visuality within heterogenous contexts of communication, production and action. For example, Askanius (2020) introduces video activism as an anchoring practice for a bundle of other activist practices such as mnemonic archiving or mobilization. Other research focuses on specific image-related practices and their embeddedness into everyday-life, e.g., “(photo-) sharing” (Murray, 2008), “camera-witnessing” (Andén-Papadopoulos, 2014), selfies-in-context (Cambre & Lavrence, 2023; Walker-Rettberg, 2014), or images in practices of intimacy (Venema & Lobinger, 2020).

Applying practice theory, the conceptual focus shifts to situated action. This shift demands the analytic inclusion of the context of use wherein images are produced, technically embedded, distributed or (re)negotiated. The notion of practice highlights routinized patterns of action (Reckwitz, 2002) which are constituted by recurring linkages of materials, meanings and competences (Shove, Pantzar, & Watson, 2012). The notion of media practices highlights the role of (technical) media within those practices. Whereas the concept of “media-related practices” (Couldry, 2012) tends to focus on media practices as a distinct class (e.g., watching TV, phoning), the concept of “media-in-practices” (Mattoni, 2020) emphasizes the kinship of very different mediated and non-mediated practices, for example differing forms of “organizing” in social movements. Accordingly, visual research can be said to either study image-related practices (e.g., photo-sharing) or images-in-practices (e.g., practices of remembering, being intimate etc.).

Research on image practices draws primarily on qualitative methods. This methodological foundation is suitable as it offers the possibility of in-depth analysis and the understanding of contextual embeddings. Accordingly, interviews, observations, or ethnographic participation are prominent approaches to delve into the visual practice of individual actors, groups, or in relation to specific events (Lehmuskallio, 2012; Luhntakallio & Meriluoto, 2022; Schreiber, 2017). Social movement and digital activism research established a rich body of literature concerning “visual protest” and image-related activist practices (Aidan, Erhart, Eslen-Zi-
ya, Jenzen, & Korkut, 2020; Doerr, Mattoni & Teune, 2013; Rovisco & Veneti, 2017; Wetzstein, 2017).

To sum up, we find profound analysis of visual representations or motifs, and the embedding of those forms in practices and everyday life. Qualitative approaches remain important for typologizing in both domains. The above methods are capable of processing small to medium sized data corpora. In addition to the established research traditions, new players stemming from the fields of digital methods (Rogers, 2021) and cultural analytics (Manovich, 2020) are concerned with digitally assisted image analysis in the realm of big data research (Drainville, 2023 within this Thematic Section). Computer vision promises to shed light on visual patterns in large samples (Niederer & Colombo, 2019; Thorsen & Astrupgaard, 2021).

Iconology usually divides the act of image interpretation into three levels: the pre-iconographic description, the iconographic analysis, and the iconological interpretation as most abstract and holistic confluence of the different elements of knowledge (Müller, 2011). In particular, computer vision is suggested to help identifying patterns on the level of preiconographic description, i.e., the primary subjects within motifs (Geise, Rössler & Kruschinski, 2016). Forms of algorithm-based sorting of images have become quite successful in terms of object recognition and formal composition (Peng & Jemmot, 2018). Recurring to Searle, the strengths of computer vision can be seen in extracting “basic perceptual features [which] are ontologically objective” (Araujo, Lock, & van de Velde, 2020, p. 242). More problems cause the identification of conventional subjects on the level of iconographic analysis, e.g., motifs from the rich history of religious allegories or political gestures and symbols.

Big data analysis in the area of social media carries an implicit kinship to practice-based research, as the corpora often rest on data extracted from “natural communication.” By revealing patterns within huge data collections, it can contribute to the mapping of the emergence of public discourse on a large scale (Waldherr, Geise, Mahrt, Katzenbach, & Nuernbergk, 2021, pp. 164–165). In the case of images, metadata or “features” (Manovich, 2020, p. 126) resulting out of basic interaction metrics (likes, reposts, comments etc.) are frequently included. Together, these data allow for basic understandings of information flows or the popularity of image trends (e.g., Mooseder, Brantner, Zamith, & Pfeffer, 2023, p. 9). At the same time, they lack a deeper understanding of images in practices; - as boyd & Crawford (2012, p. 670) put it elsewhere: “Taken out of context, data lose meaning and value.”

In what follows, we opt for a research perspective which relies on computational support, but also relates patterns of visual representations back to their communicative contexts. Image network analysis makes it possible to cluster larger amounts of images without narrowing down on predefined genres, such as memes or selfies. Additional standardized coding serves for a minimum knowledge on contextual embeddings which otherwise most often falls short in big data analysis of visuals in social media.

3 Methodological approach

In our research project on “performative publics,” we follow a relational approach, pointing to the co-constitution of public discourse by different groups of actors in social media (Lünenborg, Raetzsch, Reißmann, & Siemon, 2020). Hereby, we understand platforms as shared media spaces where heterogeneous actors play out “practices for public connection” (Raetzsch & Lünenborg, 2020) around certain issues through big and “small acts of engagement” (Picone et al., 2019). Interested in “big pictures” as well as in in-depth views, we developed a mixed-methods approach using computational, standardized and qualitative approaches (Reißmann, Siemon, Lünenborg, & Raetzsch, 2022). Starting with Twitter data, our approach systematically differentiates the contributions of societal groups such as institutionalized journalism, civil society, or private individuals through analyzing their tweet characteristics.

Exploring the role of images on the meso and macro level asks for an approach which extends practice analysis beyond individual singularity, while preserving a minimum of contextual embeddings. Therefore, we
present a methodological exploration which combines automated analysis of images with a standardized analysis of tweet practice(s) in which those images are embedded. Our ambition is to give an example for a process of research which balances two basic needs that “medium data research” asks for: reducing information to a level which standardized research can still operate with, while preserving context as much as possible. Before presenting and reflecting our exploration, we outline the background and the steps of the proposed method.

3.1 Quantitative practice profiles of public connection: A multi-method approach

Our methodological concept finds its roots in a practice-based mixed-method approach (Reißmann et al., 2022) which combines social network analysis, standardized codings of social media posts, and ethnographic case studies on influential actors (Figure 1). Point of departure are Twitter data.

Within this approach, SNA is used as a mapping tool to trace the attention that several groups of actors receive and give in different strands of discourse on Twitter. We systematically differentiate six groups of actors – journalism, non-institutionalized media, politics, civil society, science/education, and private persons (without clear affiliation to the aforementioned groups); as well as men, women and nonbinary persons. Based on the SNA, a standardized analysis for a quoted sample of tweets per actor group is carried out with the help of a codebook. Additional ethnographic case studies recontextualize the activity of selected actors beyond Twitter and social media.

Our findings from these three research steps reveal different forms of practice profiles. On the micro level: (1) individual and highly contextualized profiles for those accounts which take part in ethnographic case studies. On the meso- and macro-level: (2) group-specific profiles, aggregated on the basis of standardized actor and practice analysis; (3) Elements of Practice profiles, based on distinctive correlations in elements of practice (EoP profiles) instead of predefined actor groups. Below we show how automated image analysis can be integrated into the formation of group-specific and EoP profiles.

The quantitative practice profiles translate practice theory into standardized social media research. However, such an approach is accompanied by specific requests and challenges. Whereas ethnographic methods allow to observe the formation and dispersion of practices in an open-ended, grounded theory style, quantitative content analyses demand for pre-defined categories and

---

Figure 1: Mixed-methods approach of the overall research project and additional implementation of image network analysis

Social network analysis  
Standardized actor and practice analysis  
Qualitative case studies (interviews, ethnography)

Affiliation and Gender  
Recontextualization

Practice profiles  
Group profiles  
EoP profiles

Image network analysis

Source: Own illustration.
clearly delimited data sets. To preserve as much contextuality and cross-linkage of elements as possible, the codebook contains a wide range of categories. This gives the opportunity to keep an attentive eye for patterns on a level of descriptive quantification. However, the natural flow of practices transcending platforms and media cannot be modelled by this way of research. Nevertheless, our approach allows for systemic research on group-specific and EoP profiles.

The data corpus for the tweet analysis consists of 50 actors per societal group with max. five original tweets per actor. Thus, the total amount of tweets analyzed in each case study varies, with a potential maximum of 1500 tweets (= 50 actors × 6 groups × max. 5 tweets). Each tweet is assigned to one or more basic practices, to different topics, and to several stylistic features.

In the coding procedure, we treat images as quasi-material, meaningful objects and as one element of postings. Using a standardized method, we confine the analysis to a few characteristics (see supplementary material). As shown below (chapter 4.2), first group-specific comparisons concerning the communicative style of certain actor groups can be made due to the coded information. However, this information is rather reductive from the point of view of practice theory and visual communication research. Due to workload limits we are not able to add more depth into the codebook, for example by implementing vast batteries of potentially occurring image types. Furthermore, as visuality in social media is dynamic and diverse (Highfield & Leaver, 2016), a system of predefined categories is rather inappropriate to reveal the “natural” patterns of posted images. We therefore explore the applicability of image network analysis to examine which kinds of visualities are prevalent. In order to contextualize the findings from the automated analysis, we relate the results – distinct clusters of images – with the standardized analysis of the tweets in which those images occur.

3.2 Image network analysis: Clustering images through SNA

Today, various forms of computational methods are available for the analysis of visual data sets which most often are drawn from social media (Rogers, 2021). As an additional extension of our method, we decided to perform an image network analysis in order to identify clusters of similar visuals (Niederer & Colombo, 2019, p. 54).

In a first step, image characteristics were automatically identified for each image. Therefore, we used an image tagging approach which tags certain elements such as objects or people. For example, an image including a dog with sunglasses would get the tags “dog” and “sunglasses.” This tagging was done using the Clarifai API. Clarifai is a company in the field of Computer Vision and offers several tagging algorithms (“models”) with focus on face recognition, food, colors etc. Furthermore, it gives researchers the possibility to train a model themselves. Due to the testing character of our analysis and because there is not only one certain type of images in our dataset (such as food images or portraits), we decided to use the “general” model. Following Clarifai, this includes 11 000 different concepts including objects, topics as well as moods and should represent “a great all-purpose solution for most visual recognition needs” (Clarifai, 2022a). We got access to the Clarifai API through the open source software Memespector which was developed by Jason Chao from the Digital Methods Initiative at the University of Amsterdam (Chao, 2021). This is an easy-to-use interface which returns 20 tags for each image and a probability value of how likely a certain tag fits this image.

In a second step, an image network with the images as sources and the tags as target nodes was created with the open source software Gephi (Bastian, Heymann, & Jacomy, 2009). The probability value (see above) defines the weight of the edges. Thus, images which share similar tags are positioned closer to each other in the network.

In a third step, we performed a community detection analysis with the multi-level algorithm (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008) in order to identify image clusters and to test if and how these automatically generated clusters show coherent images. In terms of terminology, we prefer the notion “visual/image cluster” over “visual/image type.” The latter is usually the result of a sequential interpretation process, building
on multi-step, comparative ways of analysis inspired by cultural scientists like Panofsky (Grittmann, 2019, pp. 531–534). We understand unsupervised image clustering as the result of automated, algorithmic ordering principles usually based on pattern recognition on the (pre)iconographic level. Speaking of types asks for further elaboration and interpretive work.

Analyzing and publishing social media and visual data raise research-ethical questions (Heise, 2015; Wiles, Coffey, Robinson, & Heath, 2012). The visual products of our analysis are network depictions and image panels for each cluster. Following general recommendations for non-reactive research with larger data corpora, our considerations are guided by two main criteria: degree of suggested privacy and sensitivity of the data. All images were drawn from publicly accessible Twitter accounts. The recontextualization of the tweets allows us to identify speaker positions ranging from persons of public life and corporate accounts to private actors, and thus evaluate possible harms for different groups. The overall discourse frame is rather public and societal, pointing to shortcomings and needs of political action. Each image visible in this publication was assessed in terms of suggested privacy or publicity, and possible harms. Furthermore, we decided to generally anonymize depicted persons by inserting bars, as photographs in the collection eventually were taken without their knowledge (e.g., in the context of demonstrations). Additionally, intentionally low-resolution image data complicate reverse image search while offering a minimum of illustration — which we nevertheless deem essential in visual communication research.

4 Exploring the proposed method

In this chapter, we explore potentials and limitations of the outlined methodological framework using a data set of \( n = 180 \) images. The automated image tagging by Meme-spector produced a network with 10 clusters in total (modularity = 0.5). Four major clusters make up 156 out of the 180 images, displayed in green, pink, orange and purple (Figure 2). By zooming into the network and comparing the images comprised, we label the clusters as follows: “protesting people,” “nature, city, buildings without people,” “icons, symbols, graphics,” and “portraits with & without text.” Considering their composition, the crucial question is to what degree the images gathered in each cluster can be regarded as homogeneous. Suggesting unity of images turns out as the most difficult and critical point within the research process. As the clusters shall serve as a basis for subsequent recontextualization, it is important to clarify what they are standing for and if they provide enough internal consistency.

The four major clusters differ in terms of their internal homogeneity. The most consistent one is the pink cluster on “protesting people” (Figure 3). It consists of 18 images, all of them photos, most of them showing crowds in public places. Thirteen images contain clearly visible textual messages, predominantly displayed on banners and posters. In some cases, texts are inserted through post-editing. Even without detailed iconographic-iconological analysis, the dominant motifs can be subsumed to the image type

---

1 Images were downloaded for the purpose of the methodological exploration presented here. Not all images related to the data set which was created for the overall research project were available anymore (for example due to deleted accounts). Duplicates were deleted from and the respective tweets from the German pandemic Twitter hashtag #systemrelevant. The data collection started in March 2020 and ended in December 2020. The chosen sample is suitable for the purpose of methodological exploration. The rather small number of images can still be managed manually which allows us to prove automated groupings and to reflect on its benefits and pitfalls.

4.1 Image network analysis: Visual clusters of #systemrelevant

The automated image tagging by Meme-spector produced a network with 10 clusters in total (modularity = 0.5). Four major clusters make up 156 out of the 180 images, displayed in green, pink, orange and purple (Figure 2). By zooming into the network and comparing the images comprised, we label the clusters as follows: “protesting people,” “nature, city, buildings without people,” “icons, symbols, graphics,” and “portraits with & without text.”

Considering their composition, the crucial question is to what degree the images gathered in each cluster can be regarded as homogeneous. Suggesting unity of images turns out as the most difficult and critical point within the research process. As the clusters shall serve as a basis for subsequent recontextualization, it is important to clarify what they are standing for and if they provide enough internal consistency.

The four major clusters differ in terms of their internal homogeneity. The most consistent one is the pink cluster on “protesting people” (Figure 3). It consists of 18 images, all of them photos, most of them showing crowds in public places. Thirteen images contain clearly visible textual messages, predominantly displayed on banners and posters. In some cases, texts are inserted through post-editing. Even without detailed iconographic-iconological analysis, the dominant motifs can be subsumed to the image type

---
of protest and demonstration photography, or to related expressions of public solidarity. Exemplary tags are “crowd,” “motion,” “demonstration,” but also abstract ones such as “democracy,” “solidarity,” or “austerity.” Only three images do not fit into the theme of protest (No. 6, 15, 17). Notably, these three images do appear in the periphery of the cluster, not in its center.

Like the pink one, the green cluster of “nature, city, buildings without people” assembles almost only photos. Exemplary tags

![Figure 2: Clusters of images in #systemrelevant](image)

Note: Pink cluster (top middle): “protesting people” (n = 18); Green cluster (top right): “nature, city, buildings w/o people” (n = 27); Orange cluster (bottom right): “icons, symbols, graphics” (n = 57); Purple cluster (bottom left/center): “portraits w/ & w/o text” (n = 54).

Readers of the print version can find the colored figures in the online version of this text at https://www.hope.uzh.ch/scoms/article/view/j.scoms/2023.03.3883.

---

2 No. 6 does not show protesting people (though a protest-related poster of the German trade union ver.di is in the background). No. 15 resembles the visual pattern of holding up posters, but written messages reveal a different context. No. 17 is a photo of a brochure, which has a crowd of demonstrating people as a photographic element in the bottom.

are “house,” “statue,” “concrete,” or “landscape,” “countryside,” or “drink.” A basic similarity of the images is the absence of humans. In the graphic network view, the cluster splits into two halves: In the top right, natural motifs dominate, e.g., depictions of forests and parks, views of a garden. More below and on the left side, photos from public spaces appear, e.g., a monument, walls, windows, the inside of a subway. Most of the latter photos include text messages, e.g., posters and banners in front of windows. Comparing the motifs gathered in the green and the pink cluster, two differences emerge. First, motifs are much more diverse in the green cluster. Whereas one part of the images creates associations of freedom, of the worth of nature, of relaxation or beauty, the second part transports meanings of publicly visible protest in urban spaces through textual messages. Consequently, it makes little
sense to treat the different motifs as a unity on the level of meaning. Secondly, the comparison of the green and the pink cluster reveals that automated sorting \textit{may} lead to image \textit{types}, \textit{when} iconological abstraction and recognition of patterns on the level of (pre-) iconographic forms do converge. This is the case when photos of protesting people are clustered. Here, algorithmic sorting equates with a highly conventionalized symbolic form.
But it is surely not the case when the network analysis clusters all impersonal images which contain various outdoor elements.

Beyond that, the green cluster fits to demonstrate the significance of the placement of images in the network. One example is a thumbnail of a news broadcast video of German public television that is neither placed in the center of the green cluster nor in its periphery, but within the purple cluster on “portraits with & without text.” Image tags are “man,” “ball-shaped,” “universe,” and “geography” inter alia. Obviously, the two animated globes in the background of the news anchor caused the allocation to the green cluster. As it also meets characteristics of the purple cluster, it is located within those images depicting “portraits with or without text.” Therefore, it becomes clear that, although images are sorted to clusters, certain transitions and connections to different clusters become apparent depending on their placement within the image network. A thorough inspection of the proper placement of images in network graphics possibly surpasses the capabilities of methods aimed at processing larger data sets. Still, for additional qualitative purposes, this fact is essential in understanding the mixing of visual attributes and forms.

The orange cluster “icons, symbols, graphics” also assembles heterogeneous motifs. At the same time, rather high homogeneity is given on the level of visual genres: The cluster includes almost no photographic elements. It is dominated by layout and graphic design. Exemplary tags are “abstract,” “vectors,” “typography,” “graphic design,” and “illustration.” Accordingly, a meaningful basis for context-oriented concluding could be the question as to which groups of actors use rather creative visual products to participate in #systemrelevant. Yet again, sub-clusters such as “data and scientific graphics” could be separated. Diagrammatic illustrations concentrate in the bottom right corner.

The purple cluster “portraits with & without text” is the most scattered. The middle-left part is dominated by images depicting one person, or one character with no or little text. Figure 4 illustrates the heterogeneity of visual forms gathered here: from photographic portraits, to screenshots of live broadcasts, to fictive Peanut’s character Linus van Pelt, and podcast ads using the iconic portraiture of Karl Marx. Here, the tool evidently reaches its limits. The cluster mixes fictive and documentary styles, TV stills, and photos. To continue with this collection would be useless. Either manual coding or supervised tools are required to obtain more plausible results.

Other parts of this cluster combine very different visual genres and motifs alike. Only
in the lower part, at the point of transition to another (small) cluster, we find coherent images. These are predominantly share-pics showing portraits of single speakers with their textual statements/quotes. Just taking together those images alone could be a solid basis for subsequent conclusions on who uses this specific visual style of making political claims.

4.2 Recontextualizing visual clusters through practice profiles

The categorization of images is a precondition to make claims on their practical use. However, only contextual data provide a solid basis for statements on how images were mobilized within communicative acts. The results of the standardized tweet analysis (based on the project’s overall codebook) give first hints in that regard. They indicate how the given discourse on Twitter is conducted by different groups of actors both visually and textually. Table 1 shows the total amount of tweets containing one or more images, and differentiates the predominant forms of visuality.

In the case of #systemrelevant, civil society and private individuals are the groups which contribute most often through authored (text-only or visual) tweets. All other groups participate significantly less with own content. Images are included in approximately one third of the tweets originating from politics (35%), civil society (34%), and science/education (29%), and in one fifth of the journalistic tweets (19%) as well as tweets from non-institutionalized media (17%). The lowest share of images is in private individuals’ tweets (12%). Moreover, it becomes apparent that the posting of tweets that include two or more images is primarily conducted by actors from civil society and by private individuals. This practice is almost not performed by journalism, politics, and science/education.

With regard to different forms of visuality, politics and civil society appear to use more portraits than the other groups, whereas overall photographic diversity seems to be highest in civil society. Diagrammatic images are clearly the domain of scientific/educational accounts which post almost no photos. Memes are not relevant within this data set (according to the used definition, see supplementary material). This can be a hint for the overall serious character of this specific discourse, albeit many share-pics including texts and images being coded as symbolic images.

Going beyond that basic information, the next step of our method aims at recontextualizing the image clusters produced by the image network analysis. In alignment with the project’s methodological vocabulary, we attempt to abstract “EoP profiles.” Accordingly, analysis is not directed at groups of actors in the first place, but starts with “elements of practice,” here image clusters, asking for their relation with groups, anchoring practices, and modalities of public connection.

For demonstration purposes we limit this last research step to the orange and the pink cluster, for they show the highest homogeneity, in different regards. Recontextualizing the images to the tweets in which they are embedded, these clusters specify as follows:

<table>
<thead>
<tr>
<th>Table 1: Results of standardized coding of images in tweet sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td>Total amount original tweets in sample</td>
</tr>
<tr>
<td>Tweets with images</td>
</tr>
<tr>
<td>– one</td>
</tr>
<tr>
<td>– several</td>
</tr>
<tr>
<td>Tweets with photographic images</td>
</tr>
<tr>
<td>Tweets with symbolic images</td>
</tr>
<tr>
<td>Tweets with memes</td>
</tr>
<tr>
<td>Tweets with diagrammatic images</td>
</tr>
</tbody>
</table>
The tweet analysis helps to better understand how the pink cluster “protesting people” is mobilized as a visual type. Sharing visual protest is obviously the domain of civil society in #systemrelevant. Tweets do not only most often originate from this group, but sources of civil society actors are referenced the most as well. Thus, visual protest is done rather self-referentially. At the same time, accounts posting protest images inform about

| Table 2: Descriptive EoP profiles of the pink and orange cluster in absolute numbers of tweets per category |
|---|---|---|
| **Groups of actors** | **Pink cluster “protesting people” (n=18)** | **Orange cluster “icons, symbols, graphics” (n=57)** |
| Degree of organization | Corporate actors | 7 | 19 |
| | Collective actors | 6 | 10 |
| | Individual actors | 5 | 28 |
| Affiliation | Journalism | 0 | 0 |
| | Non-institutional media | 1 | 6 |
| | Politics | 3 | 3 |
| | Civil society | 11 | 19 |
| | Science / Education | 0 | 7 |
| | Private individuals | 3 | 22 |
| Gender (only coded for individual actors) | Female | 2 | 9 |
| | Male | 3 | 18 |
| | Non-binary | 0 | 0 |
| | Not identifiable | 0 | 1 |
| **Anchoring practices (multiple codes possible)** | | |
| Informing | Self-referential informing | 10 | 32 |
| | Referencing to information from others | 10 | 16 |
| Valuing | No strong valuation | 16 | 45 |
| | Strongly negative valuation | 2 | 11 |
| | Strongly positive valuation | 0 | 1 |
| Intervening | Posing demands | 10 | 8 |
| | Calls for action | 3 | 6 |
| **Modalities** | | |
| Personal experience | No personal articulation | 12 | 54 |
| | Speaking as spokesperson for others’ personal experiences | 5 | 3 |
| | Speaking as self-affected actor | 1 | 0 |
| Emotional tonality | Factual reporting | 14 | 33 |
| | Angered/disappointed/frustrated/sarcastic | 4 | 21 |
| | Delighted/humorous | 0 | 3 |
| | Offensive/hateful | 0 | 0 |
| **Sources and references (multiple codes possible)** | | |
| External links | Providing external links | 4 | 20 |
| Areas of references (= concrete references in tweet text, or linked/attached content) | Reference to Civil Society | 14 | 21 |
| | Reference to Non-Institutionalized media | 0 | 4 |
| | Reference to Journalism | 1 | 5 |
| | Reference to Science | 0 | 11 |
| | Reference to Politics | 2 | 3 |
| | Reference to Media Culture [films, series etc.] | 0 | 2 |
| | Individual reference | 3 | 8 |
their own and about others’ actions. Posing demands is the dominant form of discursive intervention. Furthermore, a rather serious and detached tweeting style can be observed. Own lived experience is revealed in only one tweet. All other tweets are written from the perspective of spokespersons or do not contain any personal articulation at all. Correspondingly, serious tonalities are very prominent. Though highlighting grievances and posing demands may convey emotional response, civil society agents rather report than mobilize affectively here. The rather low numbers of external links and concrete calls for actions foster the overall documentary style of the tweets.

The orange cluster shows both commonalities and differences compared to the pink cluster. The sharing of icons, symbols and graphics is mainly conducted by civil society members and private people alike, thus hinting at the fact that here, too, the practice of informing is based on own-group references. Consequently, sources stemming from civil society are used as reference most commonly. However, contrary to tweets with images of protesting people, the sharing of icons and symbols involves an overall more diverse group of actors. For one thing, there is no clear dominance of the civil society sector as was the case within the pink cluster. Instead, the posting and sharing of images are practiced by private individuals with the same intensity as civil society actors. For another, tweets are also posted from actors belonging to the science/education group, whereby this group was completely absent in the pink cluster. This applies to the use of references as well since, overall, the orange cluster shows a higher degree of variety. For example, as the above table shows, the orange cluster demonstrates a relatively high number of tweets that are based on references from science and academia, while there was no single scientific reference found in the pink cluster. Interestingly, in terms of modalities, the majority of tweets do not demonstrate any personal articulation, aligning with the fact that the emotional tonality of the tweets tends to be rather factual, without having a strong valuation. In fact, strikingly, there is no single tweet that seems to voice any personal experiences. Compared to the “protesting people” cluster, this cluster seems to cover a wider range in regard to emotional tonalities, for there is a good number of tweets expressing anger and frustration. This could be traced back to the variety of actors involved within this cluster overall. Especially the prevalence of private individuals might account for the fact that the tweets carry a bigger emotional load.

We only illustrate the logic of the proposed method here. The descriptive findings of the clusters’ practice profiles cannot claim significance in statistical terms. Nevertheless, as a further mapping instrument, they are informative and give orientation about the embedding of the different forms of images in practice(s). Working with larger samples and bigger clusters would make numbers and shares more reliable.

5 Discussion and conclusion

The combination of image network analysis and standardized practice analysis is a promising way for research projects which attempt to grasp the (co-)constitution of public discourse in social media through different groups of actors on an aggregated level. Our methodological exploration shows how such an analysis could look like. In conclusion, we reflect on potentials and limitations of the proposed method.

First of all, more research following a “medium data approach” should be conducted. In social media, similar kinds of images are mobilized by heterogenous actors with diverging interests, embedded in multiple practices. Thus, generalizations on the cultural meaning of certain image types become increasingly difficult. In #systemrelevant, the cluster “protesting people” is predominantly embedded in practices of visual documentary by civil society agents themselves. Here, almost no journalistic account uses this visual type for the purpose of their own posts. This does not mean that journalism or other groups never use this image type as a visual resource. They are more likely to visually cover protests in public spaces in other strands of discourse. Therefore, discourse-specific analysis which differentiates the communicative practices of different groups of actors
is suitable. The approach we presented gives a systematic basis for comparing groups and their posting activities.

Secondly, public discourse spans over different social media and platforms. A Twitter bias in much research on digital public spheres has been rightly pointed out (Lewis & Molyneux, 2018). In our project, we counter this bias with ethnographic case studies on central actors. These case studies confront our overall quantitative mapping strategies with individual practice, including their visual practice (Reißmann et al., 2023). Nevertheless, also on the basis of aggregated data, it is desirable to compare the performative making of issues on different platforms. Images related to #systemrelevant on Twitter do not necessarily coincide with those on Instagram or Facebook (Pearce et al., 2020, for the need for cross-platform analysis). The proposed method principally allows for the comparison of platform-specific image networks and their contextual embedding through the posts they are used in.

Thirdly, our exploration shows the difficulty to categorize images and to treat them as if they belong together. This is a critical point within all methods which deal with larger image corpora. The combination of image networks and contextual practice analysis calls for a higher degree of reflexivity regarding the construed character, as the ambition is to draw conclusions as to who uses certain types or clusters of images, and in which practices they are embedded. Thus, researchers should clarify to which degree a bundle of images can be treated as a unit, before making statements about its distribution among journalists, politicians, civil society etc. As seen above, an automatically produced cluster may stand for a visual type. Grouping together “protesting people” is an example for the convergence of pattern recognition on the level of (pre-)iconographic “description,” and iconological abstraction on the level of cultural meaning. The cluster – by chance – resembles an established type of visual culture and political iconography. However, other clusters in our exploration incorporate very heterogeneous forms of images and motifs. A cluster which consists of too many different visual forms and types cannot reasonably be related to the results of the standardized tweet analyses. It would remain unclear in which regard specific visual patterns relate to certain communicative practices and other tweet characteristics. We see two strategies to deal with this issue:

The first strategy is to strictly limit interpretative conclusions. We applied this strategy to conclude on the orange cluster (bottom right in Figure 2) which consists of a wide range of primarily text-based or symbolic share-pics and icons. We may draw conclusions on who uses designed visualizations in which kind of posts. In that regard, the cluster is homogenous. However, we cannot make substantial statements on the usage of specific forms and visual contents of share-pics, icons etc. This is clearly a limitation. The more general the conclusions are, the less informational value they contain. The second strategy (not applied here due to test character and small sample size) would take more advantage of the visualization of networks which follows the principle of core and periphery. Delimiting the data set to the center, or to other denser heaps within clusters may raise the probability to gather motifs and text-visual arrangements with higher internal consistency. Both strategies imply more human efforts and interpretation that follow the automated grouping. Nevertheless, we still consider a general detection model as suitable for point of departure because specific visualities (e.g., celebrity, portraits etc.) should not be prioritized in advance for the kind of research presented here.

Fourthly, working with image networks of images means working with “metapictures” (Rogers, 2021, p. 1, recurring on Mitchell). The visualization is very helpful as a mapping strategy and allows for a quasi-tangible navigation through the images. We see a huge potential for further qualitative navigation through the clusters, reflecting on close and distant relations between images, or on images which belong to one cluster but are positioned within other clusters. This navigation experience includes exciting surprises where the grouping through computer vision challenges human vision and may reveal new relations between images. At the same time, the metapicture’s role in shaping results has to be reflected as well. The image network rests on natural data and thus seem
to “mirror” real-life practice. However, it is not congruent to everyday practice of those who visually participated in #systemrelevant. They never approach the issues they feed with images from the vantage point of a network map.

Finally, the automated analysis of images raises different ethical questions. It reinvokes general questions regarding the legitimacy of processing larger data corpora which are based on non-reactive data collections. We took responsibility by assessing all published images in terms of suggested privacy or publicity, possible harms, by anonymizing faces and through providing only low-resolution image data wherever needed. Beyond that, specific – and new – challenges arise due to the implementation of commercial software. We got access to free software and are grateful for the opportunities Memespector offered. However, this software uses Clarifai API for tagging the images. Clarifai does not reveal further information on how the model was trained and which kind of tags it includes. Furthermore, it is a private company whose services are not only offered to civilian but also to economic, military, and state agencies (Clarifai, 2022b). This also applies to similar solutions provided by companies like Google or Microsoft. From an ethical point of view, the straightest path would be to rely (only?) on civil society technology with open data regimes, and which explicitly and permanently exclude uses which contradict the civilian character of social science. Against that background, we urge for an intense discussion within the field of visual communication research based on the question as to how to deal with these tensions and where to draw the boundaries regarding implicit, non-intended cooperation with commercial actors. In the end, this also touches on questions of funding, as powerful tools require substantial resources to be developed and maintained.

Acknowledgements

This paper was enabled by the funding of German Science Foundation / Deutsche Forschungsgesellschaft (DFG), project number 418375831.

Conflict of interest

The authors declare no conflict of interest.

Supplementary material

Supplementary material for this article is available online in the format provided by the authors (unedited). https://www.hope.uzh.ch/scoms/article/view/j.scoms.2024.01.3883

References


