



Selected Papers of #AoIR2021:
The 22nd Annual Conference of the
Association of Internet Researchers
Virtual Event / 13-16 Oct 2021

LOCALIZING CONTENT MODERATION: APPROACHING THE ORIENTATIONAL SPACES OF FACEBOOK GROUP ADMINS AND MODS

Dr Scott Rodgers
Birkbeck, University of London

Dr Liam McLoughlin
Birkbeck, University of London

Dr Andrea Ballatore
King's College London (from September 2021)

Dr Susan Moore
University College London

Abstract

What is content moderation? Who decides what is allowed or not on digital platforms? Emerging research suggests there is a growing diversity of answers to these questions. Content moderation is increasingly understood as practices and techniques emerging through manifold intermediaries, including platform self-regulation (e.g. Medzini, 2021), moderation algorithms (e.g. Gillespie, 2020), government policies, laws and regulations (e.g. Gorwa, 2019), an increasingly globalized industry of human moderators (e.g. Roberts, 2019), and the ways in which platform users by accident or design moderate one another (e.g. Matias, 2019; Schneider, 2021).

In our paper we address a different question: *where* is content moderation? We approach this broad question by focusing on localized social media – an area which remains under-researched (though see e.g., Miller, 2016) – summarizing early results from in-depth interviews with the content moderators of 12 place-named Facebook groups across Greater London, UK. These are groups related to named places such as neighborhoods, streets, districts, or villages. Unlike platforms such as Nextdoor, which allocate users based on their geographical location, place-named Facebook groups are created, joined, and contributed to by multiple users, built around different and often overlapping geographical scales and topics. Through a combination of automated and

Suggested Citation (APA): Rodgers, S., McLoughlin, L., Ballatore, A., and Moore, S., (2021, October). Localizing Content Moderation: Approaching the Orientational Spaces of Facebook Group Admins and Mods. Paper presented at AoIR 2021: The 22nd Annual Conference of the Association of Internet Researchers. Virtual Event: AoIR. Retrieved from <http://spir.aoir.org>.

manual web search, we identified 3,226 Greater London Facebook groups with 1000+ members (accessing public information, e.g., group name, size, description) devoted to interests and livelihoods ranging from parenting, to buy and sell, to local history. From this larger set of groups (which are being analyzed in another paper for their geodemographic relationships) we purposefully sampled 16 groups with a more general focus on a named place, seeking a mix of locations (e.g., inner and outer London), group sizes (small and large) and origins (e.g., groups initiated by individuals, businesses, municipalities).

Through the in-depth interviews, we sought to explore how moderators negotiate spatial 'orientations' between 'translocality' and 'locality'. We expected moderators would contend with Facebook as a 'translocal' space since, wherever the platform is used, it retains the same, technical functionalities for users, and is mediated through the same algorithmic and data-driven logics. The same community standards and appeals process applies, with the final global arbiter being the new Facebook Oversight Board. On the other hand, we also expected moderators would orient significantly to the localized situation surrounding the place-named Facebook group.

Four more explicit themes emerging from the interviews all related to moderators' local orientations. First, neighborhood Facebook groups are clearly venues for '*folkloric*' (see Phillips & Milner, 2017: 25-26) interactions and dynamics. For moderators, what is appropriate or inappropriate – allowed or not – is often based on tacit understandings of a 'local' culture – which might be quite different to, and can contend with, Facebook's global rules. Second, the moderators we spoke with often sought to *curate* their local group. They saw their role as stimulating contributions and conversations – often around a normative ideal of local community – just as often as it might have been about 'policing' content. Third, local moderators tend to share some perception about their *power* to support or detriment certain groups. They could see that the rules they established could for example help or hinder local businesses, or rule in, or out, certain types of political discussion. Finally, local moderators described how they faced the potential for *physical* or *place-based* harm. They and their group members usually shared the same geographical turf. We heard accounts of, for example, verbal abuse at the local pub, physical violence - even human feces left outside one moderator's home.

Our in-progress analysis of the interview data suggests that 'translocal' orientations tended to be treated more implicitly. The contingencies or conditions of Facebook as a global platform were largely taken-for-granted, or mentioned only when unusual or even extreme circumstances arose (e.g., Facebook expelling users or suspending groups). However, this implicit presence of the platform underscores, we argue, the extent to which content moderation, and social media user practices more generally, are inherently local forms of experience in which translocal platforms are materialized or actualized. As Loukissas (2019) suggests, all data are local, produced by people operating devices at particular moments, in certain situations, and oriented to various audiences or users.

This research represents one of the first systematic studies of moderation practices relating to neighborhood social media, spread across a major metropolitan region. However, we would also argue that our approach informs research on online content

moderation and platform governance more generally. Our approach to *localizing* content moderation is not one of placing moderation into pre-given geographical locations. Rather than seeing 'translocal' and 'local' as nested or fixed spatial scales, we suggest content moderation practices might be seen through the lens of relational approaches to space – spanning human geography (Marston et al., 2005; Massey, 1994), cultural anthropology (Appadurai, 1995) and philosophy (Malpas, 2012) – which see locality as a kind of *orientation* produced through situated practices. Social content moderation, understood as a media-related practice, is less a matter of studying such media practices vis-à-vis named localities, but rather studying “‘where people are’ with and through media” (Couldry and Hepp, 2017: 90–91).

The philosophical underpinning of our approach is broadly phenomenological, highlighting how what users say and do on social media is fundamentally dependent not only on the platform dynamics, but also at the same time situated experience. So, for moderators to take any stance on social media content – be it concern, worry, affection or indifference – they must first be thrown into differently-structured worlds in which social media matters. Taking the orientational spaces of content moderation more seriously allows for a far more context sensitive approach. But it also might allow scholars to better grasp and understand the limits of large-scale moderation, whether by platforms or governments, or through human or algorithmic interventions.

References

- Appadurai A (1995) The production of locality. In: Fardon R (ed.) *Counterworks: Managing the Diversity of Knowledge*. London: Routledge, pp. 208–229.
- Couldry N and Hepp A (2017) *The Mediated Construction of Reality*. Cambridge, UK: Polity Press.
- Gillespie T (2020) Content moderation, AI, and the question of scale. *Big Data & Society* 7(2): 205395172094323.
- Gorwa R (2019) What is platform governance? *Information, Communication & Society* 22(6): 854–871.
- Loukissas YA (2019) *All Data Are Local: Thinking Critically in a Data-Driven Society*. Cambridge, MA: MIT Press.
- Malpas J (2012) Putting space in place: Philosophical topography and relational geography. *Environment and Planning D: Society and Space* 30(2): 226–242.
- Marston S, Jones III JP and Woodward K (2005) Human geography without scale. *Transactions of the Institute of British Geographers* 30: 416–432.
- Massey D (1994) *Space, Place and Gender*. Minneapolis: University of Minnesota Press.
- Matias JN (2019) The Civic Labor of Volunteer Moderators Online. *Social Media + Society* 5(2): 205630511983677.
- Medzini R (2021) Enhanced self-regulation: The case of Facebook’s content governance. *New Media & Society*: 146144482198935.
- Miller D (2016) *Social Media in an English Village*. London: UCL Press.
- Roberts ST (2019) *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press.
- Phillips W and Milner, RM (2017) *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*. Cambridge, UK: Polity.

Schneider N (2021) Admins, mods, and benevolent dictators for life: The implicit feudalism of online communities. *New Media & Society*: 146144482098655.