Interdependence and Intervention: News Organizations' Contradictory Relationship with Online Comment Moderation

Anna Rantasila
Tampere University

Heli Väätäjä
Lapland University of Applied Sciences

Joel Kiskola
Tampere University

Thomas Olsson
Tampere University

Aleksi Syrjämäki
Tampere University

Mirja Ilves
Tampere University

Poika Isokoski
Tampere University

Veikko Surakka
Tampere University

Introduction

Online news comments are crucial but problematic means of audience engagement for journalistic organizations. On one hand, comments help journalists cultivating an interdependent relationship with their audiences, as comments allow the audience to communicate with journalists (Kangaspunta 2018). Comments also enable journalists to collect new information on developing stories (Masullo Chen & Pain 2017). On the other hand, the ideal of an interdependent relationship does not always match with reality. Online news comments are also a channel for uncivil discourse that can be damaging to journalists, activists, or other commenters (e.g., Frischlic et al. 2019, Winterlin et al.

2020, Wolfgang 2018a, 2018b). Uncivil commenting can also damage the brand of the publisher, which has prompted some news organizations to disable comment sections altogether (Larsson 2018).

In an ideal situation, civil commenters could discuss independently, free of outside interference. However, because of the prevalence of uncivil comments, most news organizations are compelled to intervene by moderating the comments (Gillespie 2018). Usually, moderation means prevention or deletion of unwanted messages or users (Gorwa et al. 2020, Ruckenstein & Turunen 2020).

In this paper, we examine how managers of online comments in large Finnish news organizations view online comment moderation and moderation technologies in relation to questions of inter- and independence. We suggest that while some of these issues may be solved by tweaking the design processes of comment section user interface, some may require alternative approaches to moderation altogether.

Our paper addresses the above issues via two research questions (RQ):
1) What kind of tensions the interviewees observe in online news commenting and moderation?
2) What are their opinions on the current and future technological moderation applications?

Based on these questions, we discuss the implications of the interviewees' insights for future moderation applications and practices.

Most research on online content moderation focuses on moderators or social media platforms (e.g., Caplan 2018, Gillespie 2018, Jhaver et al. 2019, Roberts 2019). Thus, this paper contributes new information by examining those who manage moderation for journalistic organizations. Moreover, this study contributes to the discussion about online content moderation by exploring how to broaden understanding of moderation beyond deletion of unwanted content and users, suggesting a move towards practices that seek to cultivate discussion (Caplan 2018, Ruckenstein & Turunen 2020).

Materials and methods

Eleven participants (9 males, 2 females; ages 33-52) were interviewed following a semi-structured procedure. The interviews were conducted in Finnish either face-to-face or via Microsoft Teams in 2019 and 2020 and audio recorded. The recordings were transcribed, and the transcriptions were further analyzed with discourse analysis.

Our sample provides a good representation of the Finnish news industry: the interviewees were employed by the six most prominent news outlets in Finland, including a major television channel, a news agency, four large newspaper publishing companies, and a software company catering to news organizations. Working in specialized middle-management positions, the interviewees influence future acquisitions and development of comment moderation systems in the media organizations. Hence, they have a considerable role in decisions that affect the online environment where news commenting takes place.

Findings and discussion

In respect to RQ1 the interviewees would like to see more meaningful engagement with their audiences. They wished for the comments to be deliberative, well-constructed, and informative. They argued that good moderation should intervene with the commenters' freedom of expression as little as possible. However, the interviewees reported frustration because the comments tend to be emotional, derogatory, and reactionary, thus requiring continuous moderation. Moreover, they saw moderation as time-consuming and outside "actual" journalistic work.

Regarding RQ2, most of the interviewees' organizations had either outsourced or looked to outsource comment moderation, and some used automated moderation. However, some interviewees were not satisfied with the outsourced or automated moderation, mostly because of a lack of contextual knowledge. This reflects the results of previous research (e.g., Caplan 2018, Gerrard 2018, Gorwa et al. 2020, Jhaver et al. 2019, Laaksonen et al. 2020, Tubaro et al. 2020).

The interviews presented three points of consideration for future moderation applications and practices, as well as research. First, there is a need to involve moderators and news organizations better in the design processes of automated moderation applications to increase understanding of the contexts where the applications will be used. Second, the interviewees suggested alternative approaches to moderation altogether. For example, replacing comments with other means of interaction was mentioned as a way to avoid uncivility. Third, to address the contradiction between freedom of expression and need for moderation, we advocate a view that would focus on rewarding civil comments instead of punishing for uncivil comments, in line with Ruckenstein & Turunen (2020). In other words, we suggest rethinking both intervention and interdependence.

References

Caplan, R. (2018) *Content or Context Moderation?* New York, NY: Data & Society Research Institute. Retrieved September 3, 2020 from: https://datasociety.net/output/content-or-context-moderation/

Masullo Chen, G. and Pain P. (2017) Normalizing Online Comments, *Journalism Practice*, 11:7, 876-892, DOI: 10.1080/17512786.2016.1205954

Frischlich, L, Boberg, S. and Quandt, T. (2019). Comment Sections as Targets of Dark Participation? Journalists' Evaluation and Moderation of Deviant User Comments, *Journalism Studies*, 20:14, 2014-2033, DOI: 10.1080/1461670X.2018.1556320

Gerrard, Y. (2018) Beyond the hashtag: Circumventing content moderation on social media, *new media & society*, 20:12, 4492–4511. DOI: 10.1177/1461444818776611

Gillespie, T. (2018) *Custodians of the Internet. Platforms, content moderation and the hidden decisions that shape social media*. New Haven: Yale University Press.

Gorwa, R, Binns, R. and Katzenbach, C. (2020) Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society* 2020 January–June: 1–15, DOI: 10.1177/2053951719897945

Jhaver, S, Birman, I, Gilbert, E. and Bruckman, A. (2019) Human-Machine Collaboration for Content Regulation: The Case of Reddit Automoderator. *ACM Trans. Comput.-Hum. Interact*. 26:5, (July 2019), 35 pages. https://doi.org/10.1145/3338243

Kangaspunta, V. (2018) Online news comments: Social network and emergent public. *The Information Society* 34:5, 275–288, https://doi.org/10.1080/01972243.2018.1497741

Laaksonen, S-M, Haapoja, J, Kinnunen, T, Nelimarkka, M. and Pöyhtäri, R. (2020) The Datafication of Hate: Expectations and Challenges in Automated Hate Speech Monitoring. *Frontiers in Big Data*, 3:3. DOI: 10.3389/fdata.2020.00003

Larsson, A. O. (2018) Assessing "The Regulars"—and Beyond, *Journalism Practice*, 12:5, 605-623, DOI: 10.1080/17512786.2017.1338149

Roberts, S. T. (2019) *Behind the Screen. Content moderation in the shadows of social media*. New Haven: Yale University Press.

Ruckenstein, M. and Turunen, L.L.M. (2020) Re-humanizing the platform: Content moderators and the logic of care. *new media & society* 22:6, 1026–1042. DOI: 10.1177/1461444819875990

Tubaro, P, Casilli, A.A. and Coville, M. (2020) The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society* 2020 January–June: 1–12, DOI: 10.1177/2053951720919776

Wintterlin, F, Schatto-Eckrodt, T, Frischlich, L, Boberg, S. and Quandt, T. (2020) How to Cope with Dark Participation: Moderation Practices in German Newsrooms, *Digital Journalism*, DOI: 10.1080/21670811.2020.1797519

Wolfgang, J.D. (2018a) Cleaning up the "Fetid Swamp", *Digital Journalism*, 6:1, 21-40, DOI: 10.1080/21670811.2017.1343090

Wolfgang, J.D. (2018b). Taming the 'trolls': How journalists negotiate the boundaries of journalism and online comments. *Journalism*, 1–18, DOI: 10.1177/1464884918762362