

# BOOK OF ABSTRACTS

21-23 November 2022

## #ERNC22

Esports Research Network  
Conference

# Sustainable Esports in the Digital Society

Jönköping, Sweden

# Understanding Twitch Esports Communities through Livestream Chat Analysis

Raine Koskimaa  
University of Jyväskylä  
[raine.koskimaa@jyu.fi](mailto:raine.koskimaa@jyu.fi)

Tanja Välisalo  
University of Jyväskylä  
[tanja.valisalo@jyu.fi](mailto:tanja.valisalo@jyu.fi)

Jari Lindroos  
University of Jyväskylä  
[jari.m.m.lindroos@jyu.fi](mailto:jari.m.m.lindroos@jyu.fi)

**Keywords:** Esports community, social sustainability, inclusion, Twitch, chat, Twitch audiences.

For the esports industry today, social sustainability has been identified as a primary challenge for the sustainability (cf. WCED 1987) of the whole industry (Nyström et al. 2022). Social and mental health, diversity, and inclusion have been identified as the most pressing issues in this area (Nyström et al. 2022.).

Audience engagement is central to the sustainability of any sport. Chat is one of the most distinctive forms of engaging audiences in esports livestreams, and it also gives viewers the possibility to participate in the livestream. Moreover, following the chat plays a significant role in experiencing esports online. The popular livestreaming platform Twitch provides multiple opportunities for audience engagement, all revolving around the chat feature. Therefore, chat plays an important role in the construction of inclusivity for audiences.

There is an existing and expanding body of research on the inclusivity of esports, though it has not always been framed in the context of sustainability (e.g., Hayday & Collison 2020; Ruvalbaca et al. 2018). These studies have also shown the existence of sexism (Nakandala et al. 2017) and racism (Gray 2016) in Twitch chat. However, the majority of the research has used fairly small amounts of data analysed with qualitative methods, with only few studies opening the way to using massive dataset and machine learning methods (e.g., Song, Park & Cha 2021; Loures et al. 2020). In order to form a comprehensive understanding of social

phenomena such as inclusion, a deeper knowledge based on large datasets of the structure of Twitch chat and its contents is needed. Who has a voice in Twitch chat?

To answer this research question, we executed a pilot study in connection to the CS:GO Majors played in May 2022. We chose this event because of the global popularity of the game and the esports surrounding it. Our aim is to analyse the chat structure and dynamics in the Majors but, furthermore, develop tools and methods of analysis to be used in other esports as well. Data in this pilot study consisted of the livestream video, and the chat data collected using the Twitch API. As the Majors were streamed on several channels, we collected chat data from three different channels: (1) PGL, the channel of the primary international livestream, (2) Pelaajat.com, a Finnish esports channel, and (3) YLE, the channel of the Finnish national public broadcasting company. Collecting data from three different channels enables us to make comparisons and perhaps detect phenomena independent of what is going on in the livestream of the game itself. Capturing the chats live provides significantly more metadata in comparison to collecting chats later, when the streams are published on the Twitch website as recordings. The live data collection also includes banned messages, which are not available later on.

We are interested in whether it is possible to discern specific viewer profiles based on viewers' use of the chat. Firstly, we analyse how the amount of chat content is divided among chat users. Additionally, Twitch chat participants have different roles based on their user account - the metadata reveals the roles of admin, moderator, and subscriber. In addition to this, user badges can reveal more detailed information. Comparison of these different types of users with the amount of content created reveals some of the basic dynamics of the chat. Secondly, we analyse the type of content created by users and aim to discern between repetitive (copy-pasta,

common phrases, emoji-only messages) and unique message content. This will reveal potential viewer profiles based on the types of chat participation that are typical for them. For identifying these types of content, we apply natural language processing methods using machine learning models. Twitch chat provides a unique challenge for these methods, as it often includes colloquial language with many specialised expressions specific to game cultures in general, a particular game, or the Twitch platform, or even a particular Twitch channel.

The tools created here will create opportunities for working with large amounts of data, such as making comparisons between different channels or doing longitudinal analysis of changes in chat structure and content. Furthermore, identifying structures and dynamics of Twitch chats

form a basis for further analysis of different communication styles, dynamic relationships between stream content and chat content, and human-in-the-loop analysis via visualisation and retrieval of notable streams, moments, and their similarities and relationships to known demographic and other covariates. On the basis of the chat dynamics modelling, we present initial analyses of how comments on nationality and gender are used in the chat as ways for discrimination and harassment. This is key to creating more inclusive chat practices and services.

## References

- Gray, K. L. (2016). “They’re just too urban”: Black gamers streaming on Twitch. In *Digital sociologies* (pp.355–368). Daniels, J., Gregory, K. & MacMillan Cottom, T. (Eds.) Bristol, UK: Policy Press
- Hayday, E.J. & Collison, H. (2020). Exploring the Contested Notion of Social Inclusion and Gender Inclusivity within eSport Spaces. *Social Inclusion* 8(3)
- Loures, T.C., Fernandes, G.L., Araújo, F.G., Martins, K.S. & Vaz-De-Melo, P.O.S. (2020) StinkyCheese: Chat-Based Model for Subscription Classification. *Proceedings of the Proceedings of ECML-PKDD 2020 ChAT Discovery Challenge*.
- Nakandala, S., Ciampaglia, G., Su, N., & Ahn, Y. Y (2017). Gendered conversation in a social game-streaming platform. AAAI Publications, *Eleventh International AAAI Conference on Web and Social Media*.
- Nyström, A.-G., McCauley, B., Macey, J., Scholz, T. M., Besombes, N., Cestino, J., Hiltcher, J., Orme, S., Rumble, R., & Törhönen, M. (2022). Current issues of sustainability in esports. *International Journal of Esports*, 1(1).
- Ruvalcaba, O., Shulze, J., Kim, A., Berzenski, S. R., & Otten, M. P. (2018). Women’s experiences in eSports: Gendered differences in peer and spectator feedback during competitive video game play. *Journal of Sport and Social Issues*, 42(4), 295–311.
- Song, H., Park, K. & Cha, M. (2021). Finding epic moments in live content through deep learning on collective decisions. *EPJ Data Sci.* 10, 43.
- WCED. (1987). *Our Common Future*. Oxford: Oxford University Press.