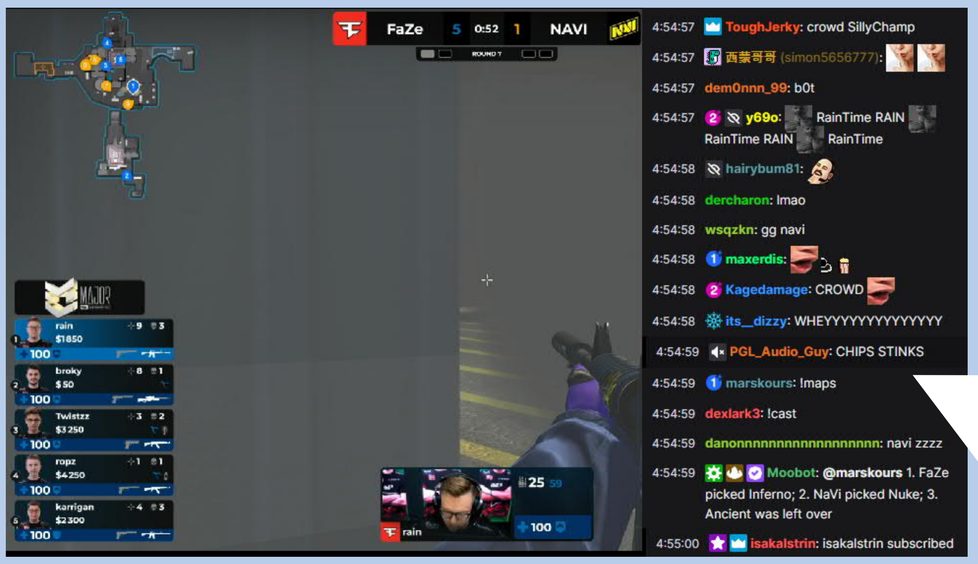
Raine Koskimaa
Department of Music, Art and Culture Studies
University of Jyväskylä
Tanja Välisalo
Faculty of Humanities and Social Sciences
University of Jyväskylä
Jari Lindroos
Faculty of Humanities and Social Sciences
University of Jyväskylä
Jaakko Peltonen
Faculty of Information Technology and Communication Sciences
Tampere University

FIN-CLARIAH

WP 3.4 Analysis tools for real-time chats in gameplay streams

One of the biggest transformations across contemporary media cultures is the proliferation of streamed audio-visual content with textual communication features. Online streams of both amateurs and professionals are changing the ways through which people engage with the world. While online streams have been studied increasingly over the past years, the employed methods we currently have are typically qualitative and have focused on limited samples. There is a lack of means for processing larger amounts of content. Thus, we need tools to study online streams quantitatively, for instance, to enable even thousands of hours in genre-profiling analyses.



Example of Twitch gameplay stream and the accompanying chat

TWITCH.TV

- Started in 2011, currently owned by Amazon
- 3 million broadcasters monthly
- 15 million daily active users, 1.4 million average concurrent users
- 27.000+ partner channels
- Categories: Games, Esports, Just Chat, Sports, Music, Creative

Twitch chat

- Participation requires logging in
- Sub-cultures based on games and genres, built around channels and casters
- Sub-culture specific communication
- Emotes, copy-pasta, cheers...
- Constantly evolving
- Connection to Alt-Right ideology

GOALS

Our goal is to develop analysis methods for real-time chat in gameplay streams.

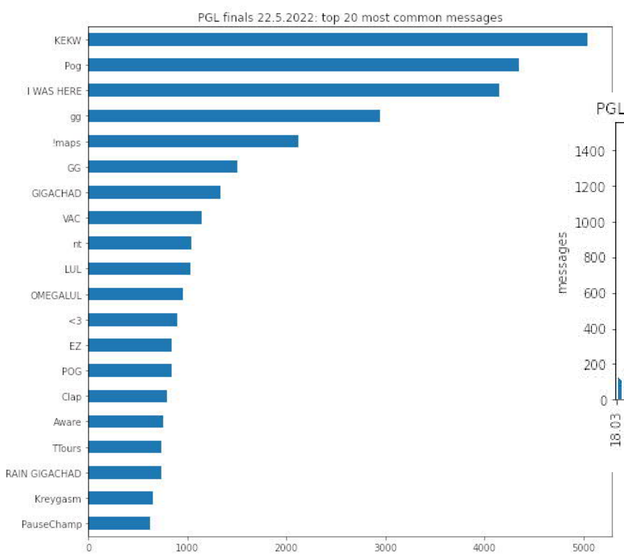
Tools for quantitative study of large masses of stream chat data

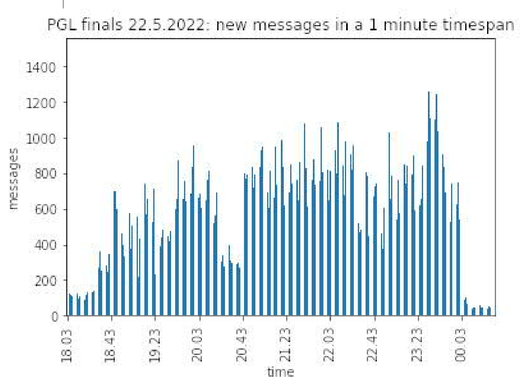
Machine learning tools to analyse stream chat structure and content

Tools for quantitative study of stream chat data. These tools enable, for example:

- analysing dynamics of stream chat: message rate, chat pace, different audience segments, how many participants
- analysing content structure of stream chat: how much 'copypasta', how much authentic data (in % of whole data), how much responses to other messages (in % of whole data), how much emojis etc.
- comparisons between different channels, and longitudinal analysis of changes in chat dynamics over time

Machine learning tools. Twitch streams offer a rich opportunity to analyze data via machine learning, and new methods will be developed for this. Tools will include modeling of differing content and communication styles via topic modeling and word embedding of chat data, extraction of repeating patterns, analysis of visual content, learning of dynamic relationships between stream content and chat content, and human-in-the-loop analysis via visualization and retrieval of notable streams, moments, and their similarities and relationships to known demographic and other covariates.





TIMELINE

Data collection pilot 5/2022 → Prototype of quantitative stream chat analysis tool by 12/2022 → Prototype of machine learning tool to analyse chat data 12/2023

TEAM

Work in this WP will be carried on by the CoE in Game Culture Studies JYU team, which has already done significant work on analyzing discussions in game streams, and the TAU machine learning team of the CoE having worked on gaming social media analytics.

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

DARIAH-FI

KIELIPANKKI The Language Bank of Finland

FIN-CLARIN